uc3m | Universidad Carlos III de Madrid

Aerospace Engineering

Bachelor Thesis 2017-2018

ESTIMATION OF TIME-RESOLVED VELOCITY FIELDS WITH MODAL ANALYSIS

Author

Ana María Sánchez Ruiz

Supervisor

Stefano Discetti

Leganés, September 2018

Abstract

Turbulence constitutes one of the still unsolved and most challenging phenomena ever covered by Fluid Mechanics. However, the state of the art methods do not provide the necessary commitment between spatial and temporal resolutions that are required for a complete comprehension of turbulent flows. This fact has boosted the development of methods relying in statistics and that are able to combine the benefits from several approaches in order to overcome such spatiotemporal limitations. Among the multiple proposals, the dynamic estimation of fluid properties based on Extended Proper Orthogonal Decomposition modes postulates as one of the most promising approaches. This method is able to estimate time-resolved turbulent velocity fields by extending the temporal correlation between time-resolved single-point velocity measurements and non-time-resolved field measurements obtained at synchronized times to the "out-of-sample" time instants. This study is based on the work performed by Discetti et al. [1] who proposed a standard truncation criterion for the filtration of the measurement signals based on the aforementioned random distribution of noise. For this purpose a parametric study is performed in order to evaluate the already proposed standard truncation threshold and two possible variants under the effects of different factors. Channel flow data generated using a DNS is used for the validation of the approach due to its large wealth of motion scales. The average estimation error is used as the figure of merit in the assessment of the filtration procedure performance and estimation quality.

Keywords: turbulent flow, noise contamination, truncation criterion, time-resolved, temporal correlation.

Index

Page

Abstract . . . iii
List of figures . . . viii
List of tables . . . ix

1 Introduction **1**
1.1 Background information . . . 1
1.2 Motivation . . . 4
1.3 Objective and structure of the project . . . 6

2 Methodology **9**
2.1 State of the art . . . 9
2.2 Dynamic estimation based on EPOD modes . . . 11
2.2.1 Mathematical fundamentals of POD . . . 11
2.2.2 Snapshot POD execution . . . 12
2.2.3 Extended POD . . . 14
2.3 Truncation criterion for the removal of uncorrelated information . . . 16
2.3.1 Some theoretical remarks on noise contamination . . . 16
2.3.2 Theoretical fundamentals of the criterion . . . 17
2.4 Potential applicability scenarios for the dynamic estimation based on EPOD modes . . . 18

3 Validation **21**
3.1 Dataset description . . . 21
3.2 Code implementation . . . 23
3.2.1 Dynamic estimation algorithm . . . 23
3.2.2 Truncation criterion implementation . . . 25
3.3 Study nomenclature . . . 25

4 Results **27**
4.1 Performance parameters . . . 27
4.2 Best case scenario . . . 29
4.3 Number of components . . . 33
4.4 Noise effects . . . 39

4.4.1 Noise effects in the absence of data filtering 39
4.4.2 Noise effects on filtered data 41
4.5 Truncation criterion 43
4.5.1 Truncation assessment for clean data 43
4.5.2 Truncation assessment under noise effects 45
4.6 Higher noise intensities 48

5 Conclusion **53**

6 Project Environment **55**
6.1 Socioeconomic impact and regulatory framework 55
6.2 Project cost 56
6.3 Project timeline 57

References **59**

Appendix **I**
PODs_and_correlations.m I
DataOrg_HW.m IV
DynamicEstimation.m V
Piv_Solution.m VII
Parametric_Study.m VIII

List of Figures

1.1 CFD Image of NASA's Hyper-X Research Vehicle at Mach 7 with Engine Operating [6]. 2
1.2 Visualization of the turbulent wake of a cylinder visualized using a laser light to illuminate tracer particles[7]. 3
1.3 Coherent structures in a mixing layer flow [9] taken from [10]. 5

2.1 Probability Density Function (pdf) of the normal or Gaussian distribution. The probabilities of intervals of values correspond to the area under the curve. 17
2.2 Basic sketch of a typical 2D particle image velocimeter system. [25] . 19

3.1 Representation of the velocity snapshots extracted from the DNS of the channel flow. In-sample slices are named with A,B,C or D, while an example of the estimated (out-of-sample) snapshots is named E. The location of the probes is set at the end (downstream) of each snapshot and uniformly distributed along the y direction between the channel wall and the centerline. Note that probes were omitted in the picture. [1] 22
3.2 Project code flowchart. 24

4.1 POD modes energy content and cumulative energy. Case 33. 29
4.2 Correlation matrices for case 33. a) Non filtered. b) 3σ filtered. . . . 30
4.3 Original snapshot, dynamic estimation with case 33 and resulting instantaneous error at an arbitrary time t. The 3σ filtered. a) Original. b) Estimation. c) Instantaneous error. 31
4.4 Correlation and determination coefficients. The 3σ rule was used for data filtering. 32
4.5 Average estimation error and average velocity correlation for case 33. 3σ rule filtered. a) Average estimation error. b) Average velocity correlation. 33
4.6 Correlation coefficient obtained for different combinations of field and single-point measurements. 34
4.7 Error estimation for all combinations of velocity components. 35
4.8 Average velocity correlation for all combinations of velocity components. 36
4.9 Ground truth velocity snapshot at a time t. 37

4.10 Snapshot reconstruction. The 3σ rule was used for data filtering. Left: Case 33. Right: Case 12. 37
4.11 Instantaneous error reconstruction. The 3σ rule was used for data filtering. Left: Case 33. Right: Case 12. 38
4.12 Average estimation error at x=0.9h. Non filtered data. 39
4.13 Average estimation error for several noise levels. Case 33 with no data filtering. a) No noise. b) 3% noise. c) 5% noise. d) 7% noise. . 40
4.14 Average estimation error at x=0.9h. All cases were filtered with the 3σ rule. 41
4.15 Average estimation error for different level noises. Case 33 filtered with the 3σ rule. a) No noise. b) 3% noise. c) 5% noise. d) 7% noise. 42
4.16 Case 33. Correlation coefficient for different truncation criteria in the absence of noise effects. 43
4.17 Case 33. Average estimation error for different truncation criteria in the absence of noise. a) No filter. b) 2σ filtered. c) 3σ filtered. d) 4σ filtered. 44
4.18 Average estimation error at x=0.9h for different truncation criteria. . 45
4.19 Case 33. Correlation coefficient for different truncation criteria. 7% noise addition. 46
4.20 Case 33 with a 7% of noise addition. Average estimation error for different truncation criteria. a) No filter. b) 2σ filtered. c) 3σ filtered. d) 4σ filtered. 47
4.21 Average estimation error at x=0.9h. No noise. 48
4.22 Correlation coefficient. Case 33 filtered with 3σ rule. 49
4.23 Average estimation error profile at x=0.9h. Case 33 filtered using 3σ rule. 50
4.24 Average estimation error for case 33 filtered with 3σ rule. 51

6.1 Project Gannt chart. 58

List of Tables

3.1 Nomenclature of the studied cases 26

6.1 Total project costs 57

Chapter 1

Introduction

1.1 Background information

Unsteadiness, irregularity and randomness are the three main characteristics present in the motion of any turbulent flow. It is for this very reason that the most often adjective associated to turbulence phenomena is chaotic. Another aspect that greatly contributes to the complexity of turbulent flow problems is the large number of spatial and temporal scales posed by such events [2].

The dynamics behind the motion of turbulent flows have been assiduously studied and there is a continuous upwelling of new experimental and numerical techniques trying to throw some light on the inherent unpredictability of this kind of flow. Nevertheless, the innate characteristics of turbulence make of its measurement and simulation an extremely complicated task. For a better understanding of this last statement, the main limitations and challenges associated to the current research status of turbulence are detailed in the following.

Numerical approaches

The well-known Navier Stokes equations, already available at the end of the 18^{th} century, can completely describe the relationship between the different properties of a moving fluid, either laminar or turbulent [3]. However the huge complexity linked to their analytical solution forced mathematicians and engineers to make further assumptions and simplifications in order to obtain a solvable system. All those simplifications consisted in the use of dynamic models which represented the mayor source of error, especially when dealing with turbulent flows. It was not until the late years of the 20^{th} century that the advances in the computational speed of computers allowed to solve them using different methods such as finite difference, spectral methods, finite volume... This fact gave birth to Computational Fluid Dynamics (CFD) [4, 5], which aims to numerically analyze and solve the motion of fluid flows. Figure 1.1 shows an example of the application of CFD for the simulation of the flow around the wing of NASA's Hyper-X Research Vehicle.

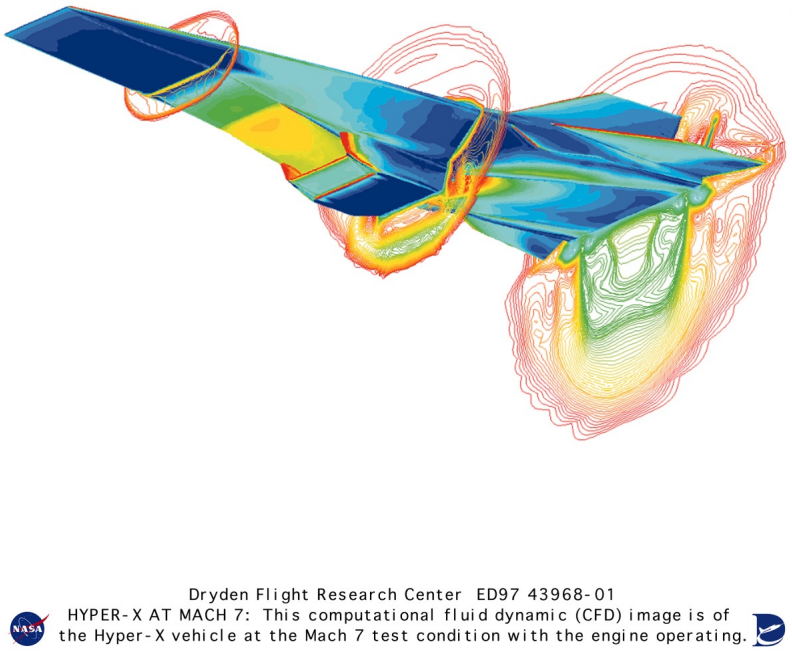

Figure 1.1: CFD Image of NASA's Hyper-X Research Vehicle at Mach 7 with Engine Operating [6].

The computing power reached today has remarkably augmented the discretization capabilities applied to fluid mechanics through the use of increasingly refined grids, thus allowing fairly satisfactory levels of accuracy. One of the most well-known examples of numerical approaches in the resolution of flow problems is the direct numerical simulation (DNS). This kind of CFD simulation considers the whole range of spatial and temporal scales present in turbulent flows, thus disregarding any turbulence model.

A full description of turbulence can only be achieved once the full velocity vector field, the pressure and temperature are known and resolved in the continuous time. These demanding conditions, i.e. the great wealth of motion scales observed in turbulent flows and the need to know their continuous temporal evolution considerably increases the computing costs associated to their simulation ($\sim Re^3$). The practical applicability of DNS simulations is then limited to the range of low-medium Reynolds numbers in the turbulent regime, while leaving the analysis of higher Reynolds number flow cases in the hands of statistics-based methods such as Large Eddy Simulations (LES), Reynolds-Averaged Navier-Stokes (RANS) and Probability Density Function (PDF) methods, which must necessarily be supplemented with appropriate turbulence models in order to deal with the closure problem [2].

Experimental approaches Following the same tendency, tremendous advances were achieved during the last decades of the 20^{th} century as a result of the great improvements of electronics, measuring sensors, lasers and the integration of optics [5]. The very first attempts of flow visualization were based on the use of fluid markers such as small particles or dyes that were added at certain selected locations in order reveal the flow path and streamlines. The reader can find illustrative examples about this topic such as the one appreciated in figure 1.2 in [7].

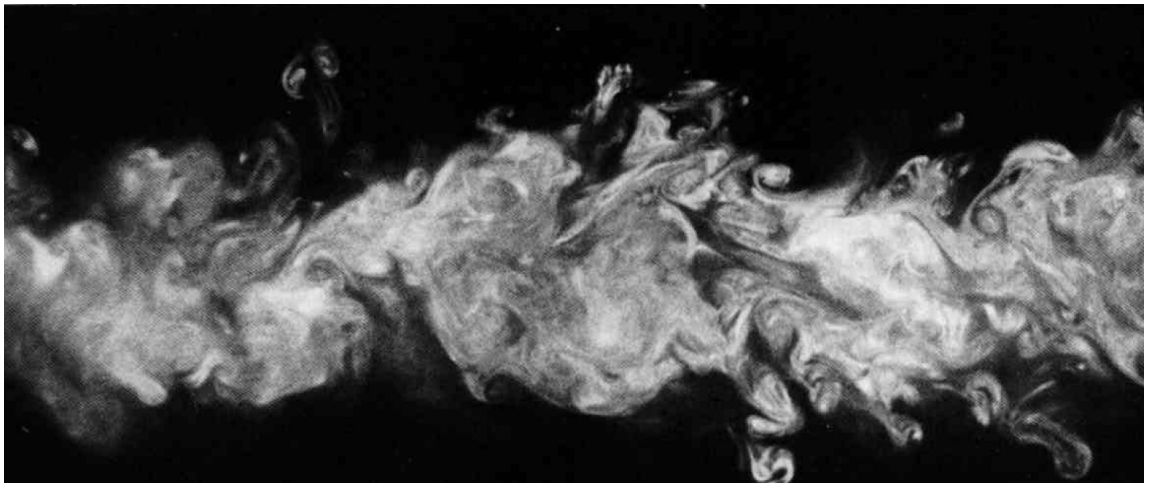

Figure 1.2: Visualization of the turbulent wake of a cylinder visualized using a laser light to illuminate tracer particles[7].

All those former visualization techniques allowed the obtaining of information that was mostly qualitative, while the responsibility of acquiring accurate quantitative data of the flow had to rely on intrusive single-point measuring methods such as hot-wire anemometry or Laser-Doppler anemometry. The great advances witnessed in the field of imaging technology and processing boosted the development of non-intrusive flow visualization techniques such as the Particle Image Velocimetry, which is capable of reproducing a 2D or even 3D both qualitative and quantitative description of the flow field[8].

However and again, the turbulent fluctuations of the velocity field vary locally from one point to another with length scales covering a wide and continuous spectrum. Thus, the accurate measurement of turbulence requires the ability to respond to that wide range of fluctuation intensities, from the smallest motion scales, whose kinetic energy is instantaneously dissipated into heat, up to the largest scales, which involve the generation of a certain net kinetic energy. Unfortunately, the state of the art sensors used for experimental measurements are just able to capture part of the whole ensemble of motion scales of turbulence and their measurement natural frequencies lie far from the fast dynamic response related to the flow fluctuations with the highest Reynolds numbers. Another issue that contributes to the complexity of

turbulence measurement is the presence of electronic noise of anemometers, which must be accounted for and somehow compensated [3].

The here depicted scenario shows that despite the important benchmarks achieved in the research of turbulence, there is still a a long way to go in the definition of tractable models for the description of engineering turbulent flows.

1.2 Motivation

Experience has demonstrated that turbulence governs the behavior of the majority of flow cases observed both in nature as well as in technology. The water in a waterfall, the flows used in turbomachinery and propulsive systems, the wind moving around airplanes and vehicles, the smoke from a chimney and also the blood running through the human veins are just a few examples of many turbulent flows present in our everyday surroundings. The main objective behind bringing these examples is to show the wide variety of circumstances in which turbulent flows occur and how diverse their demonstrations can be. It is very clear that turbulence represents a standard feature to deal with and one of the most important phenomena ever covered by fluid mechanics. That is why it appeals to the urgent need to achieve the sufficient description level to allow a complete comprehension of its behavior[3]. Turbulence is thus a complicated phenomena which can be easily observed but extremely difficult to interpret, understand and explain and the same time [9].

On the other hand, a certain coherence is found in the chaotic behavior of turbulence. Large-scale undulations, often referred to as "coherent structures" can be found in the motion of turbulent flows. Coherent turbulent structures are simple formations in which turbulent flows are usually classified and that can be used to understand their complex dynamics. They contain the major part of the flow net kinetic energy and that is why the comprehension about the formation, evolution, interaction and collapse of these large-scale structures can help to understand the turbulence phenomena. Besides that, their life period must be long enough to allow their study by means of time-averaged statistics methods that allow the dynamic estimation of these coherent structures. Figure 1.3 shows an example of the so called "coherent structures" or also "turbulent organized motion".

The already mentioned excessive complexity related to its dynamics imposes large challenging conditions to the spatial and temporal resolutions required for a reliable study of strongly turbulent flows. In the previous section the main limitations of the two main turbulence research fields were reviewed. It was mentioned how field measurement methods such as PIV allowed the illustration of instantaneous velocity fields, providing both qualitative and quantitative data with high spatial resolution. Notwithstanding the above, the still existing shortage in the temporal resolution of those techniques together with the costs associated to such innovative hardware resources discards its applicability to medium-high Reynolds number flows.

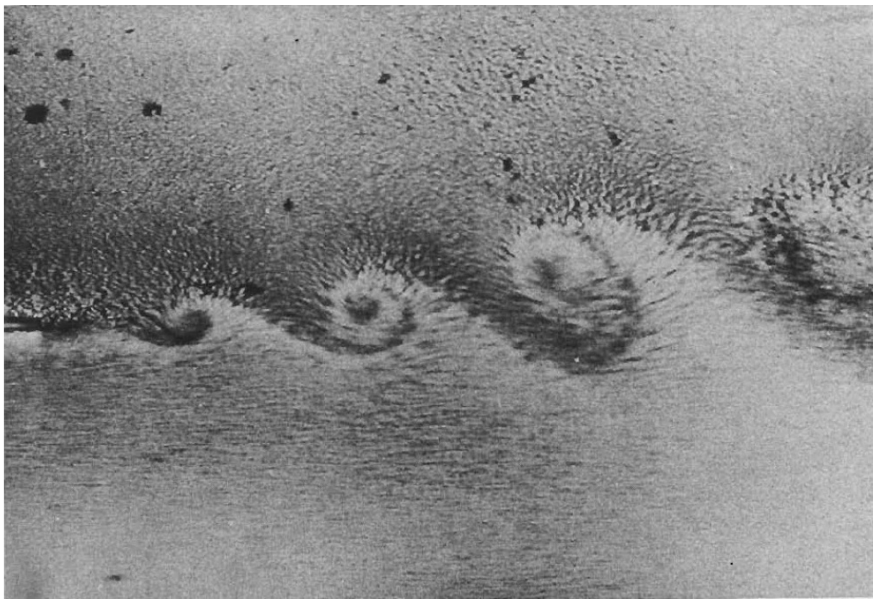

Figure 1.3: Coherent structures in a mixing layer flow [9] taken from [10].

Contrarily, single-point measurements of the flow field acquired with techniques such as hot-wire anemometry do make possible the detailed capturing of accurate quantitative data caused by small scale and high-frequency perturbations. The disadvantage is that one can not achieve satisfactory spatial resolutions since the number of individual probes required may be so large that the measurement becomes no longer non-intrusive. Thus the instantaneous description of the flow velocity field based on this method is out of reach.

In the light of this scenario in which the strengths of each individual measurement method do not compensate for their weaknesses, the idea of integrating flow data extracted with different approaches in order to take advantage of their individual potentials becomes particularly attractive. For that, the validation of powerful statistical tools such as the dynamic estimation of fluid properties is required. An attempt to answer the appealing need of disclosing the full dynamics of turbulent flows departs from the previous concept. Several methods have been proposed for the estimation of time-resolved velocity fields using non-time-resolved flow data. These methods are based on the use of stochastic estimation and dynamic models in order to identify the most relevant flow structures, and compute the out-of-sample data. For recent examples of the application of these methods refer to [11, 1, 12].

To conclude, it has been stated how the nowadays existing limitations for the simultaneous obtaining of high spatial and temporal resolutions in the description of turbulence can be potentially solved with the use of capable statistical methods such as the dynamic estimation. This opens the way to further developments of the method that may contribute to its validation against different factors such as the presence of uncorrelated information or random noise, and hence to improve the refinement of its resulting output.

1.3 Objective and structure of the project

In view of the imperative need for acquiring a complete comprehension of the turbulence phenomena, the exploiting of the dynamic estimation of fluid properties represents one of the most promising approaches that may overcome the large spatial and temporal resolution limits of the state of the art flow measurement methods. On the other hand, the extended proper orthogonal decomposition (EPOD)[13] provides a powerful tool for the dynamic estimation of time-resolved velocity fields using non-time resolved velocity measurements and a set of fast pointwise probes [12].

The aim of this study is to contribute to the validation of such technique through the assessment of a truncation criterion for the removal of uncorrelated information. The filtration of uncorrelated data in the dynamic estimation based on EPOD modes can potentially improve the accuracy of the results and the robustness of the method. The introduction of a universal filtering criterion was explored by Discetti et al. in [1] as a proof of concept. In this study, it is intended to validate the use of this standard criterion under different conditions. To that end, this work analyses the impact of several parameters such as the presence of noise effects or the number of velocity components used upon the performance of the dynamic estimation with Extended POD modes.

In order to clearly set the way forward, the project is structured in six parts. Background information necessary for the understanding of the state of the art in the comprehension of turbulent flows is provided in chapter 1. To this end, both experimental and numerical research fields are addressed and the main benchmarks and limitations associated to each of them are highlighted. After the illustration of the current turbulence problem scenario the motivation behind the project is presented. To finish the introduction to the work, the project objective and structure is briefly summarized to provide a quick overview of the problem and the strategy behind the proposed solution.

Chapter 2 details the methodology applied for the development of an algorithm that enables the dynamic estimation of time-resolved velocity fields using time-resolved single-point measurements and non-time-resolved field measurements. First the current state of the art is depicted in order to remark the need for continuous further refinements of promising early-stage techniques such as the dynamic estimation of flow properties. Then, the mathematical reasoning behind the Proper Orthogonal Decomposition is provided to continue with the specifications needed for implementation of the extended POD approach. Also the truncation criterion fundamentals and noise issues are discussed in this chapter.

The validation of the technique is shown in chapter 3 by means of its application to a given test case. The dataset and code implementation are described. Moreover some remarks regarding the nomenclature used for the representation of the results can be found at the last part of the chapter. Finally, the results obtained from the parametric study are collected and discussed in chapter 4, while chapter 5 covers

the main conclusions extracted.

The last part of the project, chapter 6, addresses the socio-economical environment and regular framework related to this study. Besides that, the project cost and organizational timeline is detailed.

Finally, the cited bibliography is listed and the discussed algorithm is added in an annex at the end of the document.

Chapter 2

Methodology

The present section explains the different steps and procedures to be followed for the dynamic estimation of time-resolved velocity fields using extended POD modes. On the other hand, the theoretical background behind the filtration criterion for its implementation and testing under a variety of noise levels is also detailed here. To put the reader in context, first the state of the art of the question is addressed. Later a brief reasoning of the mathematic fundamentals of the Proper Orthogonal Decomposition is provided to continue with the dynamic estimation based on Extended POD. Lastly, the truncation criterion to be implemented is discussed and the procedure for the introduction of noise and assessment of its effects upon the test case is presented.

2.1 State of the art

The potential behind the dynamic estimation and the promise of bridging the knowledge gap with reference to the instantaneous organization of turbulence has given place to numerous studies and attempts of validating and refining such technique.

Long gone are the days when the information corresponding to only one time instant was used to estimate the flow fields [14]. Such dynamic estimations relied on convective times for the performing of the estimations and were known as single-time-delay estimations due to the time step used between the reference data and the intended estimation.

Further approaches stand up for the operation of linear stochastic estimation considering several convective time scales which enable to reduce the amount of noise measured by the probes [15]. In the case of facing conditional and unconditional sources which present different features (either they represent different physical quantities or grids) the importance of filtering the non-correlated part of the estimations becomes even more appealing. In [16] the modeling of turbulence in a high-speed jet was implemented by applying the linear stochastic estimation to spectral data and

using the coherence spectra as a filter. An assessment study on the ruggedness of multi-time LSE-POD methods is found in [17].

Based on previous similar researches such as the ones found in [18] and [12], in 2013 Tu et al. [11] went a step ahead into the removal of noise in dynamic estimations and used a dynamic model of the flow field identified based on an initial set of estimations. Then they added a Kalman smoother for a further refinement of the results. Another interesting and innovative proposal is the one made by Baars and Tinney [19]. They took advantage of higher order coherences in the spectral domain in order to study the mutual interference between higher- and lower-energetic turbulent structures.

The present thesis bases the dynamic estimations performed on the Extended Proper Orthogonal Decomposition (EPOD) scheme. In 2003 the research carried out by Borée et al.[13] presented the concept of extended POD modes and their advantageous properties. This approach states that the temporal correlation between several synchronized signals can be assessed using extended POD modes and thus it provides a tool for the study of both temporal and spatial interactions between coherent structures. For instance, in an analysis regarding multiple synchronized velocity measurements, these extended POD modes are conceived to be the projection the velocity snapshots of one signal onto the temporal basis corresponding to the other measurement.

More recently, Hosseini et al.[20] suggested some improvements for this technique and simplified the correlation of the EPOD scheme to the temporal domain. For this purpose, the authors based the projection of the extended modes on the correlation between both temporal POD coefficients from the snapshots and probe data. Hosseini et al.[20] also claimed that this method was self-tunning since it did not consider uncorrelated modes.

The advantages associated to every approach mentioned so far are limited to the study of shedding-dominated flows. In the presence of wealthier frequency scales the achievement of an optimal truncation of the reconstructed signal gets more complex and the noise contamination becomes a critical feature to face. Prove of that can be seen in [1], where Discetti et al. extended the work done by Hosseini et al in [20], and proposed a truncation method for the removal of noise in flow field estimations. The data filtering developed was based on the random distribution of the non correlated part of the signal and also the complexity related to the choice of an optimum time-span for dynamic estimations in turbulent flows was discussed. The proposed truncation of the correlation matrix between measurements showed satisfactory filtering results and also opened the way to further explorations of a refined filtering criteria.

2.2 Dynamic estimation based on EPOD modes

2.2.1 Mathematical fundamentals of POD

In general, the Proper Orthogonal Decomposition (POD) can be defined as a process that provides a base for the modal decomposition of a set of signals. Its beauty resides in the fact that the generated base is completely data-driven without the need of any initial assumption, and that it allows to identify the dominant structures in a turbulent flow while avoiding the so feared nonlinearities [21].

The basic idea behind the application of POD to the analysis of turbulent velocity fields is the consideration that, speaking in the least mean square terms, the maximum projections onto the velocity field correspond to the coherent structures of the turbulent flow[22]. This means that the orthogonal basis formed by the first r POD modes is the one providing the minimum norm for the sum-squared error of the actual velocity values and their projections. This condition is reflected by the L2-norm in equation 2.1.

$$min\big(\sum_{k=0}^{l} \|\underline{u_k} - \mathbb{P}_r \underline{u_k}\|\big), \tag{2.1}$$

where $\mathbb{P}_r$ is the projector operator. The main consequence of this condition is that the modes obtained by means of POD are naturally ordered according to its energy contribution to the flow field. In the previous equation, $\underline{u_k}$ represents the fluctuating part of the velocity field at a given instant of time. In fact, any velocity used in this study refers to turbulent velocity unless any other definition is specified. The turbulent velocity fields are obtained after subtraction of the mean velocity field from the complete measured samples and are often referred to as "snapshots".

The condition in equation 2.1 can be solved by means of an eigenvalue problem for which its nullspace is equal to either the spatial or the temporal correlation tensor, depending on whether the approach followed is the classical POD or the Snapshot POD respectively. For simplicity, just one component $u(x,t)$ of the velocity field is used. Then, the turbulent velocity field at any time instant t $u(\underline{x},t)$ can be decomposed as follows

$$u(\underline{x},t) = \sum_{n=1}^{\infty} a^{(n)}(t)\Phi^{(n)}(\underline{x}), \tag{2.2}$$

where $a^{(n)}(t)$ is the n^{th} time coefficient and $\Phi^{(n)}(\underline{x})$ is the n^{th} spatial mode [22]. Moreover, two important properties of the decomposition must be highlighted:

- Orthonormal character of the eigenfunction basis. This property is represented by the integral domain:

$$\left(\Phi^{(n)}(\underline{x}), \Phi^{(m)}(\underline{x})\right) = \delta_{nm}. \quad where\ \delta_{nm} \begin{cases} = 1\ if\ n = m \\ \neq 1\, if\, n \neq m \end{cases} \tag{2.3}$$

- The temporal coefficients of the decomposition are independent from each other.

$$\langle a^{(n)} a^{(m)} \rangle = \lambda^{(n)} \delta_{nm}, \tag{2.4}$$

  where $\lambda^{(n)}$ is the n^{th} eigenvalue and the operation $\langle\rangle$ represents the ensemble average.

For efficiency reasons the Snapshot POD proposed by Sirovich [23] is applied here.

2.2.2 Snapshot POD execution

Starting from an ensemble of velocity field snapshots, each velocity field sample corresponds to a given time instant and is composed by a given number of point measurements. This information can be rearranged into a data matrix $\underline{\underline{u}}(n_t \times n_p)$ with t rows corresponding to the number of snapshots available and a number of columns n_p equal to the amount of points composing each of the field samples. According to expression 2.2 , such velocity data matrix can be decomposed in singular values[1] as follows:

$$\underline{\underline{U}}_{(n_t \times n_p)} = \underline{\underline{\Psi}}_{(n_t \times n_t)} \underline{\underline{\Sigma}}_{(n_t \times n_t)} \underline{\underline{\Phi}}_{(n_t \times n_p)} \tag{2.5}$$

where $\underline{\underline{\Psi}}$ is the time decomposition matrix of the turbulent velocity field $\underline{\underline{U}}$ and $\underline{\underline{\Phi}}$ constitutes the spatial basis for such decomposition. It must be mentioned that for simplicity, it has been assumed that the number of available samples is smaller than the number of point measurements per sample, i.e. $n_t < n_p$, which is a common feature in the case of, for instance, Particle Image Velocimetry experiments. Similarly, the columns of the temporal basis and the rows of the spatial decomposition matrix retain the temporal and spatial modes of $\underline{\underline{U}}$ respectively. On the other hand, $\underline{\underline{\Sigma}}$ is a diagonal matrix formed by the singular values of the fluctuating field, in descending order, that represent the energetic contribution of each mode to the fluctuating velocity matrix $\underline{\underline{U}}$.

[1] In this decomposition it has been assumed that all data in the snapshot matrix are equally grid-weighted.

Similarly, the aforementioned singular value decomposition can be applied for the analysis of velocity data obtained using a ensemble of single-point measurements. The sampling ratio offered by the majority of the state of the art single-point measurement methods, such as hot-wire anemometers, is considerably larger than the one offered by field measurement methods like Particle Image Velocimetry (PIV). Further details about the possible application scenarios for dynamic estimation, including the combination of PIV and hot-wire signals, is addressed in the last section of this chapter. Here, assume a sampling ratio $q > 1$, i.e. being the sampling rate of the fast measurements q times faster than the one corresponding to the field samples above. The approach followed here is based on the construction of "virtual" fast probes treated as to be located at downstream locations and the spatial information assigned to each of them comes from the time-delayed measurements. This procedure was successfully applied by Hosseini et al.[20] for the estimation of the velocity field in the quasi-periodic turbulent wake of a pyramid. Such authors constructed a series of virtual sensors, based on Taylor's Hypothesis, in order to correlate the historical readings of pressure sensors and the velocity field samples associated to the same time instants.

Hence, considering a number s of synchronized single-point measurements with a sampling rate q, the total number of velocity samples recorded during the same n_t time instants of field measurements above is $n_{pr} = s \times q \times n_t$, leading to a number of $n_{tt} = s \times q$ single-point measurements per velocity snapshot acquired. The n_{tt} "virtual" probe samples assigned to each sample field can be reshaped into one-dimensional row vectors and then collected into a data matrix $\underline{\underline{U}}_{pr}$. Therefore $\underline{\underline{U}}_{pr}$ is composed by the same amount of time instants (rows) as the field data matrix and a number of n_{tt} points (columns).

Finally the POD snapshot decomposition is applied to the virtual probes data matrix which has n_t rows, (the same amount of time instants as the field snapshot matrix), formed by the n_{tt} points measured in the time range between field samples. And again, it is assumed that $n_t < n_{tt}$.

$$\underline{\underline{U}}_{pr\ (n_t \times n_{tt})} = \underline{\underline{\Psi}}_{pr\,(n_t \times n_t)}\, \underline{\underline{\Sigma}}_{pr\,(n_t \times n_t)}\, \underline{\underline{\Phi}}_{pr\,(n_t \times n_{tt})} \tag{2.6}$$

The interpretation of equation 2.6 is the same as for the field data matrix and shall not be repeated here. With regard to the selection of the time interval used to determine the number of virtual probes assigned to each field snapshot the widespread practice is to relate this choice to convective processes in the turbulent flow using the average convective velocity.

The decomposition strategy explained within this section allows the reconstruction of both the field and probe snapshot vectors at any time t according to expressions 2.7 and 2.8 respectively.

$$\underline{u}(t) \approx \sum_{i=1}^{n_m} \psi_i(t)\, \sigma_i\, \underline{\phi}_i \tag{2.7}$$

$$\underline{u}_{pr}(t) \approx \sum_{i=1}^{n_m} \psi_{pr,i}(t)\, \sigma_{pr,i}\, \underline{\phi}_{pr,i} \tag{2.8}$$

where n_m is the number of modes in each reconstruction.

2.2.3 Extended POD

The next step in the strategy used in this work is the implementation of the Extended POD. The aim is to increase the number of field samples available through the dynamic estimation of the out-of-sample velocity fields based on the correlation between the in-sample field and single-point measurements. For a simpler comprehension of the process to be developed a brief enumeration of the steps to follow is found below.

1. Assessing the mutual correlation between the modes of the probe measurements and the modes of the field samples in the time domain.
2. Once this correlation is known, it can be used to determine the corresponding POD time coefficients at those time instants for which initially there was no available field sample data.
3. Reconstruction of the out-of-sample velocity fields with the previously estimated POD modes coefficients.

Correlation assessment

For the estimation of the temporal correlation between field and probe measurements first the extended POD modes of the field samples must be defined. According to the work done by Borée in [13]:

$$\underline{\underline{\Psi}}_{pr}^{T}\, \underline{\underline{U}} = \underbrace{\underline{\underline{\Sigma}}_e\, \underline{\underline{\Phi}}_e}_{EPOD} = \underline{\underline{\Psi}}_{pr}^{T}\, \underline{\underline{\Psi}}\, \underline{\underline{\Sigma}}\, \underline{\underline{\Phi}} = \underline{\underline{\Xi}}\, \underline{\underline{\Sigma}}\, \underline{\underline{\Phi}} \tag{2.9}$$

In the previous relation the subscript e is used to indicate the reference to the extended POD modes. Moreover, it can be clearly seen that the determination of any of the extended POD modes relies on the combination of all the snapshot POD modes. Notice also the definition of the sought temporal correlation matrix Ξ, where

the information about the mutual interaction of the field and probes temporal modes is collected.

$$\underline{\underline{\Xi}} = \underline{\underline{\Psi}}_{pr}^{T} \, \underline{\underline{\Psi}} \tag{2.10}$$

Out-of-sample time coefficients estimation

The just estimated correlation matrix contains information about how the single-point measurements modes correlate with the field samples modes at the in-sample synchronized time instant. Assuming working conditions in which the scale of the data sample used to estimate this correlation is sufficiently large so that statistical convergence can be reached, the following approximation can be made

$$\underline{\underline{\Xi}} = \underline{\underline{\Psi}}_{pr}^{T} \, \underline{\underline{\Psi}} \approx \underline{\psi}_{pr}^{+\,T} \, \underline{\psi}_{DYN}^{+} \tag{2.11}$$

The expression 2.11 states that once the correlation matrix between probe and field measurements $\underline{\underline{\Xi}}$ is known, it can be used at the same time as an estimate of the dyadic product between the time coefficient of the probe sample at a generic time instant $\underline{\psi}_{DYN}^{+}(t^{+})$ and the time coefficient of the field snapshot $\underline{\underline{\psi}}_{DYN}^{+}$ at any out-of-sample time instant.

$$\underline{\psi}_{DYN}^{+}(t^{+}) \approx \underline{\psi}_{pr}^{+}(t^{+}) \, \underline{\underline{\Psi}}_{pr}^{T} \, \underline{\underline{\Psi}} \approx \underline{\psi}_{pr}^{+}(t^{+}) \, \underline{\underline{\Xi}} \tag{2.12}$$

There is still one unknown to be resolved in the former relation: the obtaining of $\underline{\psi}_{pr}^{+}(t^{+})$. However this is just a mere formality recalling the expression 2.8 and solving for the desired quantity. This step is indicated below and shows the possibility of estimating the vector of time coefficients $\underline{\psi}_{DYN}^{+}(t^{+})$ departing from a probe sample acquired at a generic time instant (t^{+}) and the use of the probe snapshot modes computed with the help of equation 2.8.

$$\underline{\psi}_{pr}^{+}(t) = \underline{u}_{pr}^{+}(t^{+}) \, \underline{\underline{\Phi}}_{pr}^{T} \, \underline{\underline{\Sigma}}_{pr}^{-1} \tag{2.13}$$

Out-of-sample field snapshots reconstruction

The very last step leads to the desired estimation of velocity field at the initially non available time instants. The aforementioned computation is performed using the POD modes for the field snapshots according to equation 2.5

$$\underline{u}^{+}_{DYN}(t^{+}) \approx \sum_{i=1}^{n_m} \psi^{+}_{DYN,i}(t^{+})\,\sigma_i\,\underline{\phi}_i \tag{2.14}$$

The procedure followed so far provides a means to achieve the main goal remarked at the beginning of the section 2.2.3: matching the sample frequency of the field snapshots with that of the fast "virtual" probes. This is achieved through the increase in the sampling frequency of the first one by a factor q using the explained extended POD modes approach. The result is a new dataset with a total of $q \times n_t$ field snapshot samples.

2.3 Truncation criterion for the removal of uncorrelated information

The difficulties of dealing with turbulent flows when trying to dynamically estimate the velocity field arises from their associated richness of motion scales which considerably affects the temporal relationship between the independent measurements. Thus, in order to obtain the spatial and temporal evolution of the those turbulent structures with the highest energy content it is critical to objectively estimate the correlation between velocity signals throughout the flow domain and, for that, this study also assesses the use of a truncation criterion based on the work developed by Discetti et al.[1]

2.3.1 Some theoretical remarks on noise contamination

The data used for this study was a synthetic dataset generated using a Direct Numerical Simulation. A detailed description of the dataset is provided in section 3.1. A numerically generated dataset compiles the best conditions for a dynamic estimation since it provides a noise-free data and besides it is neither affected by features related to the limited spatial resolution. However, in the real case scenario experimental velocity measurement methods such as PIV or hot-wire anemometry suffer from noise contamination.

Discetti et al. in [1] argued that correlation does not self-vanish in every sense of the word but the non-correlated part spreads as noise in the dynamic estimation. In such case, the orthonormal base formed by the columns and rows of the correlation matrix Ξ may contain higher order uncorrelated modes that are actually accounted for into the dynamic estimation.

Two main types of errors found in experimental measurements are the bias and random errors, being the first one out of the scope of this study. Besides that, the random error often seems to be the prevailing contribution and this is why the

assessment of the truncation criterion is focused on the mitigation of its associated effects. The starting point for the filtering process is the consideration of spectrally white random error as it is shown in the next section.

2.3.2 Theoretical fundamentals of the criterion

Due to the statistical background of this approach the accuracy of the results is largely influenced by the quality of the reference dataset upon which the dynamic estimation is performed. It is at this point where the need for a filtering of the results arises.

The probability density function of a given random variable (PDF) shows the probability of the random variable to take the values within a particular range. Figure 2.1 shows the probability density function (PDF) of a Gaussian distribution with zero mean and standard deviation σ. The percentages correspond to the likelihood of the given random variable to fall within the different value intervals which are expressed as a function of the standard deviation σ.

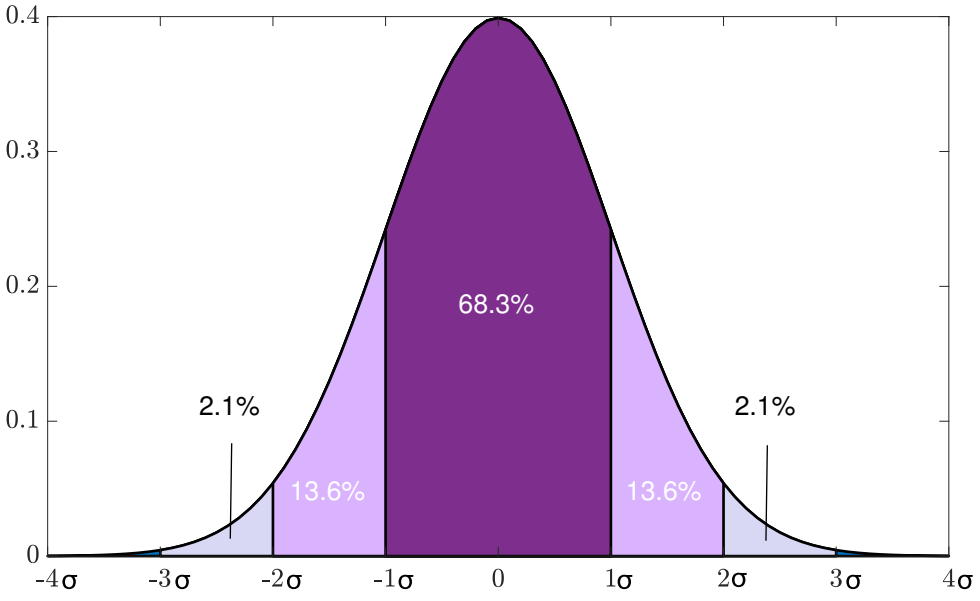

Figure 2.1: Probability Density Function (pdf) of the normal or Gaussian distribution. The probabilities of intervals of values correspond to the area under the curve.

The problem to be faced when a certain probe mode i does not correlate with a certain mode j of the field snapshots, is that this correlation does not self-vanish but its contribution is randomly spread among all the field modes. This results in a

random contamination that decreases the potential accuracy of the estimation. The aim is then to remove the randomly distributed non-correlated contributions from the correlated part of the signal.

The implementation of the filtering criterion departs from the assumption of a Gaussian error. Speaking in terms of probability, this means that the random error follows a normal distribution with zero mean and standard deviation σ. Exploiting this assumption, the temporal correlation matrix Ξ is said to be composed of randomly distributed values and a truncation error was implemented based on the inherent properties of the Gaussian distribution. Following this reasoning, the data who showed correlations within a certain mσ interval, where m is an integer multiple of the standard deviation, were eliminated and hence their noisy contribution to the dynamic estimation.

2.4 Potential applicability scenarios for the dynamic estimation based on EPOD modes

Before continuing with the technical details associated to the validation of the of proposed solution, it is convenient to provide the reader with some examples of the potential applicability scenarios for the dynamic estimation of fluid properties. One of the most extended applications that greatly benefits from the advantages associated to such approach is the combination of both PIV and hot-wire anemometry data. In order to facilitate a better comprehension of the reasons behind this data matching and the use of EPOD modes, one must go through their working principles and understand their respective potentials and limitations.

Particle Image Velocimetry

Particle Image Velocimetry (PIV) is a non-intrusive velocity measurement method. The classical PIV is known as 2D or planar PIV and its resulting output comes in the form of a planar description of the velocity field, also called snapshots. PIV technique involves tracing particles that are homogeneously seeded into the studied flow. The flow is then illuminated with a double-pulsed laser and the light scattered by the tracing particles during the first and second laser pulses is recorded by a digital imaging system onto two different image frames. Finally the flow velocity is derived from the spatial correlation analysis performed between the position of the tracer particles recorded onto two consecutive image frames. A basic sketch of a typical 2D PIV setup is shown in figure 2.2.

Since first conceived, Particle Image Velocimetry has manifested itself to be the most representative method for the estimation of velocity fields in experimental fluid mechanics, constituting one of the most powerful and polyvalent tools in the research of turbulence [24]. It has already been stated that the three-dimensional inherent character of turbulence advocates for the unavoidable necessity of a full

3D space resolution in order to obtain a complete comprehension of coherent flow structures. This fact boosted the development of different PIV-based measuring techniques towards a higher dimensionality with important improvements of the achievable spatial resolution. Some of the most noteworthy variants of PIV such as the Holographic PIV or the Tomographic PIV are able to provide 3D descriptions of the flow field[8].

However, the main drawback of PIV and even its most modern variants is their limited temporal resolution. This disadvantage relies in the fact that despite the higher pulse-frequencies achieved by even the highest performing lasers available nowadays, the scanning frequencies feasible still fall well short of the characteristic time scales observed in higher Reynolds number flows. Hence, these methods are not able to capture the fluid dynamics in presence of the wealth of scales associated to higher Reynolds number turbulent flows and are limited to the analysis of low- to medium-Reynolds number flows.

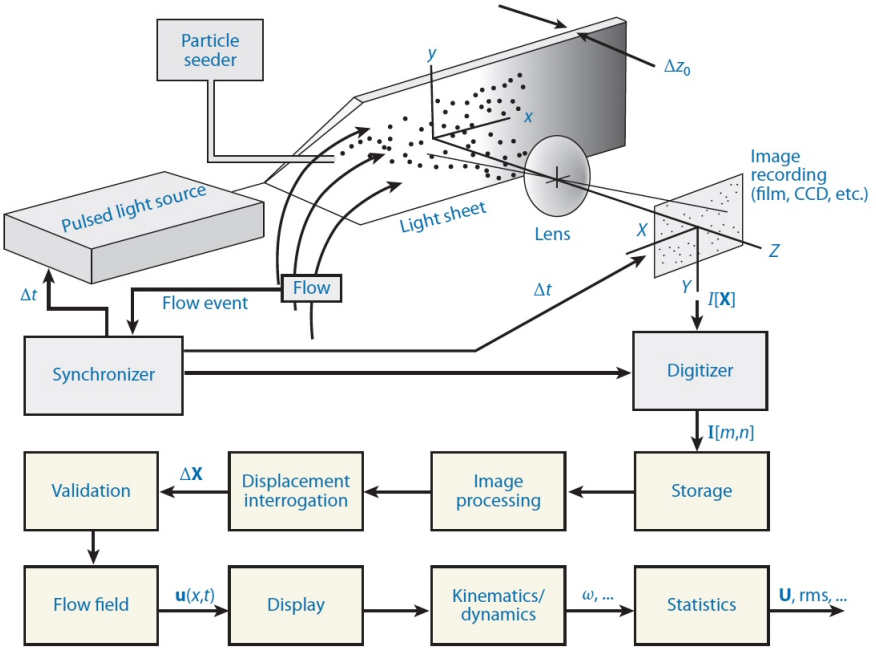

Figure 2.2: Basic sketch of a typical 2D particle image velocimeter system. [25]

Hot-wire anemometry

The basic principle behind this method relates the electrical resistance of a sensor to its temperature. With the purpose of obtaining the heat transfer between the measuring device and the surrounding flow the hot-wire is heated by an electrical current and exposed to the moving flow. Then, the electrical signal produced by the hot-wire anemometer is statistically processed in order to obtain the flow velocity. This can be done by using either analogical or digital systems. [26]

Hot-wire anemometry constitutes a high performance approach when it is intended to study turbulent perturbations whose dynamic response is especially high, i.e. the sought velocities vary at very high rates and their distorting effects want to be avoided.

It seems obvious why this method is still so popular in the investigation of turbulent flows with a wide range of turbulence scales. Unfortunately, the main handicap of single-point techniques like this one is that they are unable to provide the high spatial resolution necessary for the visualization of the instantaneous flow organization, indispensable condition for a detailed description of the coherent structures associated to turbulent flows. Further information and details about this velocity measurement method can be found in [27].

In this context, the description of the turbulent velocity field with both satisfactory spatial and temporal resolutions can be achieved by means of the combination of data obtained with PIV and hot-wire experimental techniques. The exploiting of the elevated spatial resolution of field measurements together with the high temporal response capabilities of single-point techniques can thus be achieved by means of statistically-based approaches such as the dynamic estimation used in this work. However, this is just an example of the potential applications of the dynamic estimation, which can be used to evaluate the correlation of fluid properties in general like, for example, pressure and velocity measurements.

Chapter 3

Validation

The third section concerns the development of an algorithm for the validation of EPOD approach together with a truncation criterion. First a deep characterization of the working data is provided. In the particular case of this study a synthetic data case is used to validate the truncation criterion and assess the effects of both the number of velocity components and noise upon the effectiveness of the aforementioned technique. Afterwards, the algorithm developed for the EPOD approach and the truncation criterion implementation is described. The code was designed using Matlab platform and it is attached as an annex at the end of this document in case of further interest.

3.1 Dataset description

In order to continue with the work done by Discetti et al.[1] and, at the same time, facilitate the comparison of results, the same synthetic dataset used is chosen to be the test case here. The following paragraphs provide a briefing about the details of this dataset. A more complete description is found in [1].

The interesting features for the selection of a synthetic case is that DNS provide the most beneficial scenario for a "clean" assessment of the correlation between field and single-point measurements since there is no noise or spatial resolution limitations. The flow motion through a simple rectangular channel is simulated using a Direct Numerical Simulation from the Johns Hopkins Turbulence Databases [28]. The simulation time is equal to the channel stream-wise length and the flow bulk velocity is equal to one. A sketch of the flow channel is included below for more clarity.

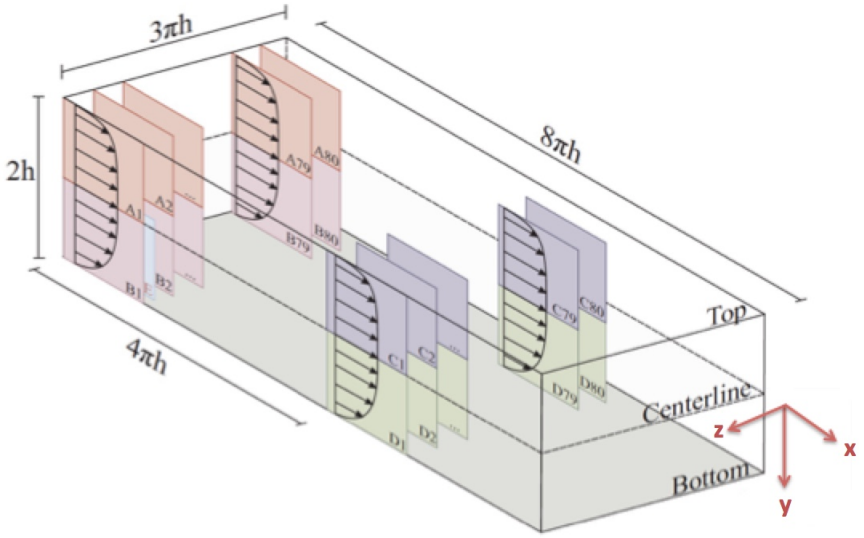

Figure 3.1: Representation of the velocity snapshots extracted from the DNS of the channel flow. In-sample slices are named with A,B,C or D, while an example of the estimated (out-of-sample) snapshots is named E. The location of the probes is set at the end (downstream) of each snapshot and uniformly distributed along the y direction between the channel wall and the centerline. Note that probes were omitted in the picture. [1]

2D PIV snapshots were obtained each one covering a spatial region of $h \times h$ in the x-y plane, where h is the height of the channel measured from the centerline, and establishing a time span between them equal to the convective time. The result is a maximum number of 24 snapshot samples per location and simulation which obviously is far from providing the necessary amount of data in order to ensure statistical convergence. Thus, the authors decided to take snapshots at 80 different locations along the span-wise flow direction at both the center and half of the channel stream-wise length. This measurement sequences were performed at the top and bottom parts of the channel thus covering the whole channel height. Hence a richer-sampled dataset with a total of $n_t = 7680$ $(24 \times 80 \times 4)$ field snapshots was acquired by taking advantage of the symmetry of the problem and assuming that the statistical properties are the same at any point of the overall flow domain. Each field snapshot is composed by $n_p = 7744$ points over a squared domain (88×88).

On the other hand, 3D virtual probe measurements were obtained at 11 single-point locations, uniformly distributed between 0.05-0.95h coordinates in the wall-normal direction (see the reference coordinate system in figure 3.1, in which the origin is taken at the centerline and the positive wall-normal direction points towards the bottom of the channel), immediately down-stream of each of the PIV snapshots previously collected. The sampling rate of the probes was arbitrarily chosen to be

704 times faster than the PIV. Recalling that the time delay between snapshots was set to be equal to one convective time, this means that a number of $704 \times 11 = 7744$ measurements are assigned to each field sample.

Lastly and following the same procedure, both probe and field samples were acquired at out-of-sample locations, i.e. at locations different from the ones used for the just estimated POD modes but at a time-step 5 times wider than before. The probe measurements at those "out-of-sample" locations (an example is the green region named E in figure 3.1) were used for the implementation of the dynamic estimation while the PIV slices were kept as a ground truth to discuss the results.

3.2 Code implementation

3.2.1 Dynamic estimation algorithm

The dataset collected from the DNS simulation was assembled into two matrices containing the data corresponding to the field snapshots and the virtual probes. The very first step into the data analysis is to properly isolate the turbulent velocity fluctuations according to Reynolds decomposition, and for that the mean velocity fields for both PIV and probe measurements were calculated and subtracted from the corresponding complete velocity fields. Afterwards these matrices were used to estimate the snapshot POD modes and the virtual probes modes using an economy size singular value decomposition in Matlab. From equations 2.5 and 2.6 and using the dataset of the test case, all this adds up in:

$$\underset{=(7680\times n_p)}{\underline{U}} = \underset{=(7680\times 7680)}{\Psi}\underset{=(7680\times 7680)}{\Sigma}\underset{=(7680\times n_p)}{\Phi} \tag{3.1}$$

$$\underset{=pr\ (7680\times n_{tt})}{\underline{U}} = \underset{=pr\,(7680\times 7680)}{\Psi}\ \underset{=pr\,(7680\times 7680)}{\Sigma}\ \underset{=pr\,(7680\times n_{tt})}{\Phi} \tag{3.2}$$

One of the factors whose effects upon dynamic estimation was to be studied is the number of components measured by both PIV and virtual probes. For this purpose 2 and 3 component PIV measurements were considered, while for the case of the fast probes 1,2 and 3 component sampling were used. This scenario leaves place for the study of 6 different combinations of PIV-probe measurements. Following this reasoning, the number of points n_p in equation 3.1 depends on the number of components considered for the PIV snapshots, being equal to 2×7744 in the case of 2-components PIV and to 3×7744 for the full three-components PIV. The same applies for the virtual probes decomposition, being here $n_{tt} = 1 \times 7744$ for the case of one single component probe measurement. Once both decompositions were performed the next step was to use the estimated temporal modes matrices in order to obtain the correlation matrix Ξ, according to equation 2.11, for the 6

combinations of PIV-probes components. Both POD decompositions as well as the computation of the correlation matrices were acquired by means of the algorithm *PODs_and_correlations.m* that can be found at the attached at the Appendix.

Next, the time coefficients for the out-of-sample snapshots are obtained using the already computed correlation matrices and the time coefficients of the probe measurements obtained at those same locations. Finally, the desired PIV velocity fields are estimated by projection of the PIV POD modes onto the just calculated non-in-sample time coefficients. All those operations were performed following the procedure described in equations 2.12-2.14 with the help of the code named *DynamicEstimation.m* also available at the end of this document. In figure 3.2 the reader is able to find an explicative flowchart of the project code.

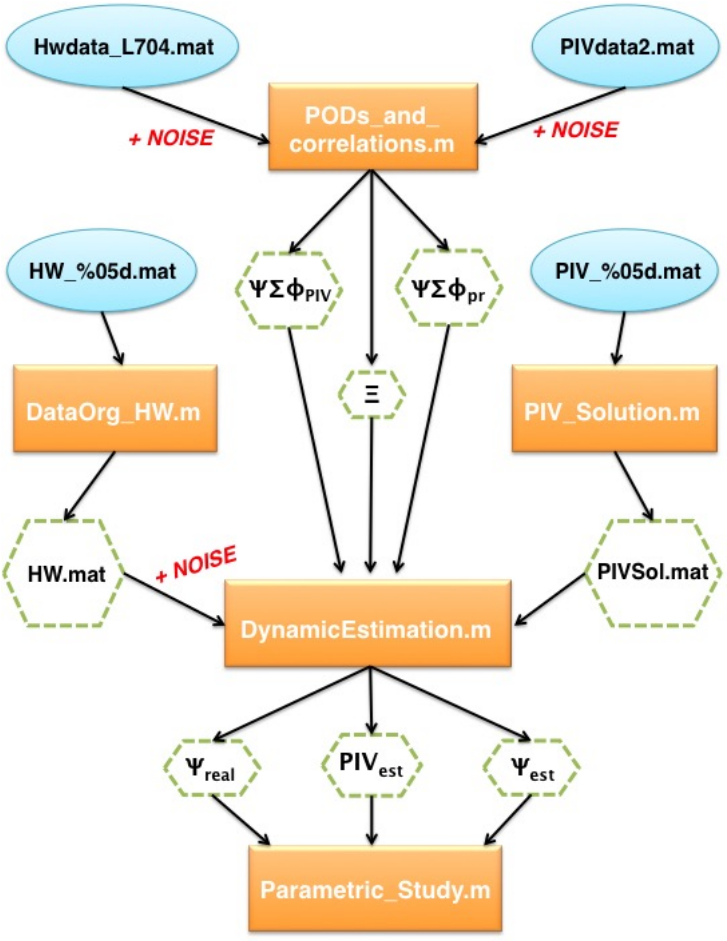

Figure 3.2: Project code flowchart.

3.2.2 Truncation criterion implementation

First the assumption made in section 2.3.2 is checked. Effectively, every row and column of the correlation matrix has a norm close to 1 and zero mean. Thus the assumption is valid and the computation of the standard deviation $\sigma = \frac{1}{\sqrt[2]{n_t}}$ is straight.

The density probability function of the error distribution is assumed to be similar to the one found in figure 2.1. As stated in the previous chapter, the percentages correspond to the probability of noise signals (vertical axis) taking correlation values Ξ_{ij} (horizontal axis) between the indicated intervals, which are expressed in terms of the standard deviation σ for more clarity. The simplicity behind this probabilistic distribution facilitates the implementation of a truncation criterion for the removal of non-correlated contributions. The work done by Discetti et al.[1] established a standard threshold for the truncation using the 3σ rule proposed by Pukelsheim[29], who showed that 99.7% of the data under a normal distribution are collected within 3 times the standard deviation of the mean. This truncation condition is shown in the equation below.

$$-3\sigma \leq \Xi_{ij} \leq +3\sigma \rightarrow \Xi_{ij} = 0, \quad i,j = 1,...,n_t \tag{3.3}$$

For the present study this truncation threshold is assessed using several multiples of the standard deviation: 2σ, 3σ (for comparison with former works), 4σ. Also non-filtered cases are considered for comparison of the results obtained after data refinement. The testing of its efficiency under several noise levels is performed by first introducing a random signal, i.e. noise, into both the PIV and probe reference measurements. The intensity of the added noise ranges from 3-25% of the observed fluctuations and its effect upon the filtering criterion and the number of field measurement and probe velocity components used for the dynamic estimation is studied. Moreover, the logic of the problem requires that this same noise signal is applied to the out-of-sample single-point measurements used for the computation of the EPOD modes and, therefore, previously to the dynamic estimation.

Regarding the implementation of the truncation criteria within the algorithm environment, it was included in the code developed for the dynamic estimations *DynamicEstimation.m*, while the different noise levels were applied to the initial reference data used to compute the POD modes and the out-of-sample point measurements in the codes named *PODs_and_correlations* and *DataOrg_HW.m* respectively.

3.3 Study nomenclature

For clarity and simplicity, some nomenclature guidelines applied in the presentation of the results obtained in this study are specified here.

The number of velocity components contained in both PIV and single-point measurements are indicated by means of two-digit number. The first digit corresponds to the number of velocity components measured in the single-point signal and the second integer indicates the velocity components captured in the field snapshots. Following this reasoning, the references used for the six different combinations of field snapshots and point measurements considered here are detailed in table 3.1.

NOMENCLATURE		
Case nomenclature	**Point-measurement velocity components**	**Field measurements velocity components**
Case 12	1	2
Case 13	1	3
Case 22	2	2
Case 23	2	3
Case 32	3	2
Case 33	3	3

Table 3.1: Nomenclature of the studied cases

Furthermore, for the analysis of noise effects the hardness of the filtration criterion applied to each case is indicated in terms of the typical deviation associated to the assumed normal distribution of noise. According to that, the 3 filtering levels considered 2σ, 3σ and 4σ indicate that correlation data within two-,three- and four-standard deviations from the correlation matrix Ξ were neglected, and so for the dynamic estimation.

Chapter 4

Results

4.1 Performance parameters

Previous to the immersion in the study of the results it is convenient to provide a clear definition of the parameters here applied to assess the effect that different factors such as the number of components considered for the estimation, the presence of noise or the filtering sharpness used can have upon the performance of the dynamic estimation technique.

- ***Squared correlation coefficient.***

  For the comparison of the dynamically-estimated time coefficients $\psi_{i,DYN}(t)$ with the real time coefficients $\psi_{i,real}(t)$ the squared correlation coefficient of the modes time coefficients was used. It is defined as follows:

$$R^2(i) = \frac{\langle(\psi_{i,DYN}(t)\ \psi_{i,real}(t))\rangle^2}{\langle(\psi_{i,DYN}(t))^2\rangle\langle(\psi_{i,real}(t))^2\rangle} \tag{4.1}$$

  Note that the angular brackets $\langle\rangle$ represent the ensemble averaging.

- ***Determination coefficient.***

  An alternative and more conservative way of assessing the performance of the dynamic estimation of the modes time coefficients is the determination coefficient. However, the use of the correlation coefficient was generally preferred in this study.

$$R^2_{det}(i) = 1 - \frac{\langle SS_{res} \rangle}{\langle SS_{tot} \rangle} \tag{4.2}$$

$$SS_{res} = \sum_{i=1}^{n_m} \left(\psi_{i,real}(t) - \psi_{i,DYN}(t)\right)^2 \tag{4.3}$$

$$SS_{tot} = \sum_{i=1}^{n_m} \left(\psi_{i,real}(t) - \overline{\psi}_{real}(t)\right)^2 \tag{4.4}$$

- ***Average estimation error***.

  The squared mean error in the velocity estimation is computed as:

$$\overline{e}^2(x) = \sum_{j=1}^{j=n_t} \left[u_{est}(x_i, t_j) - u_{real}(x_i, t_j)\right]^2 / n_t \tag{4.5}$$

  For the sake of simplicity, the simple mean error is used instead of the squared one.

$$\overline{Err}(x) = \sqrt{\overline{e}^2(x)} \tag{4.6}$$

- ***Velocity correlation***.

  In this case both the standard deviations of the real and estimated cases are used to obtain the non-dimensional form of this parameter.

$$\overline{v}_{corr} = \frac{\left(u_{est}(x_i, t_j)\, u_{real}(x_i, t_j)\right)/n_t}{\sigma_{u,est}\, \sigma_{u,real}} \tag{4.7}$$

- ***Mode energy contribution and cumulative energy***.

  They are computed using the eigenvalues of the POD decomposition, knowing that $\lambda_i = \sigma_i^2$.

$$Energy\ contribution = \lambda_n / \sum_{i=1}^{n_m} \lambda_i \tag{4.8}$$

$$Cumulative\ energy = \sum_{i=1}^{n} \lambda_i / \sum_{i=1}^{n_m} \lambda_i \tag{4.9}$$

4.2 Best case scenario

When dealing with the best case scenario it refers to the consideration of the most robust combination of measurement signals. In the context of this study this definition corresponds to case in which the correlation between field and single point measurements is performed under the best statistical convergence conditions, i.e. taking into account the maximum amount of data available. Taking into account this requirement, the case 33 in the absence of noise was chosen to be analyzed since it accounted for the correlation between the three-dimensional velocity measurements from both the field snapshots and the ensemble of virtual probes.

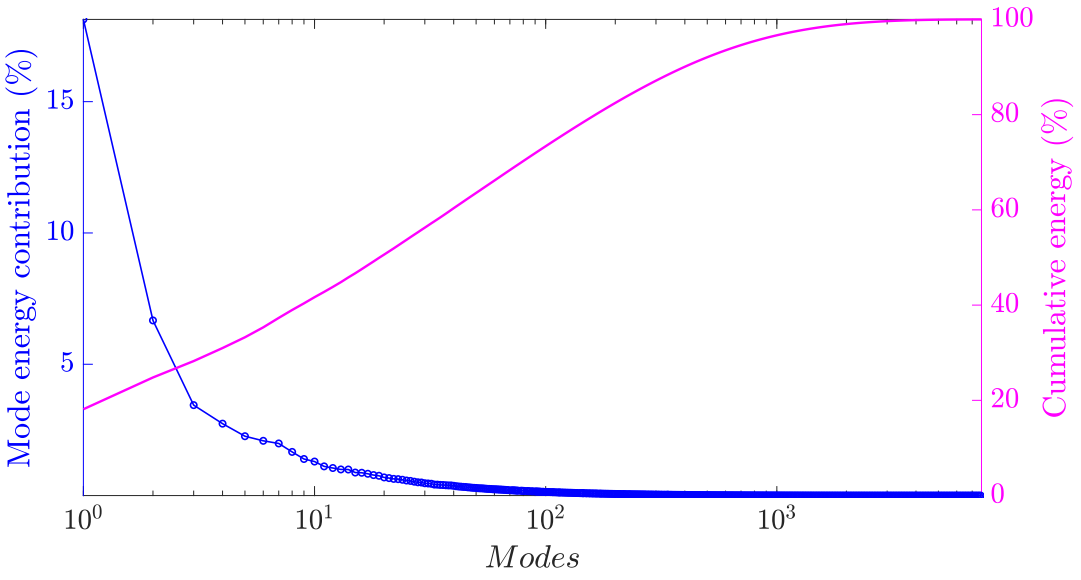

Figure 4.1: POD modes energy content and cumulative energy. Case 33.

Figure 4.1 plots the energy contribution of each mode. This is computed as the ratio between each eigenvalue, i.e. each of the squared singular values obtained from the POD ($\lambda_i = \sigma_i^2$), and the total summation of the POD modes eigenvalues. Besides, the cumulative energy content of the modes is shown. In section 2.2.1 it was demonstrated that one of the main advantages associated to the use of the POD is that the modes obtained are naturally ordered according to their energy content. Thus the energy contribution tends to vanish as the number of modes increases, being the first modes those with the largest motion scales. This fact is verified by looking at figure 4.1 which depicts the energy distribution between the POD modes extracted from the 2D field snapshots with information about three velocity components.

Also the cumulative energy of the modes is depicted here. The cumulative energy in the figure shows how the DNS test case studied is not a shedding-dominated flow

but it presents a much higher spectral richness, requiring the contribution of about 400 different modes to reach 90% of the total turbulent flow energy.

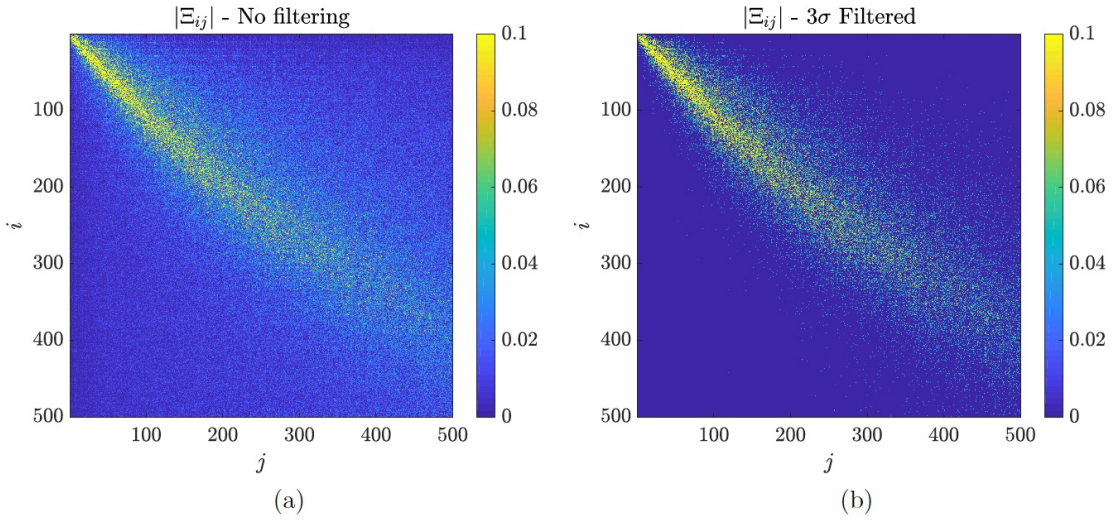

Figure 4.2: Correlation matrices for case 33. a) Non filtered. b) 3σ filtered.

Attending to the most relevant modes, the correlation matrices for both unfiltered and 3σ filtered cases are shown in figure 4.2 for the 500 most energetic modes. The diagonal dominance behavior found in both cases is highlighted when the standard filtering rule proposed by [1] is applied, thus eliminating the major part ($\approx 99.7\%$) of the uncorrelated signal spectrum. These results justify and support the need for a filtration criterion that refines the estimated velocity quantities. Therefore, any result presented henceforth counts on the standard filtration rule unless other conditions are specified.

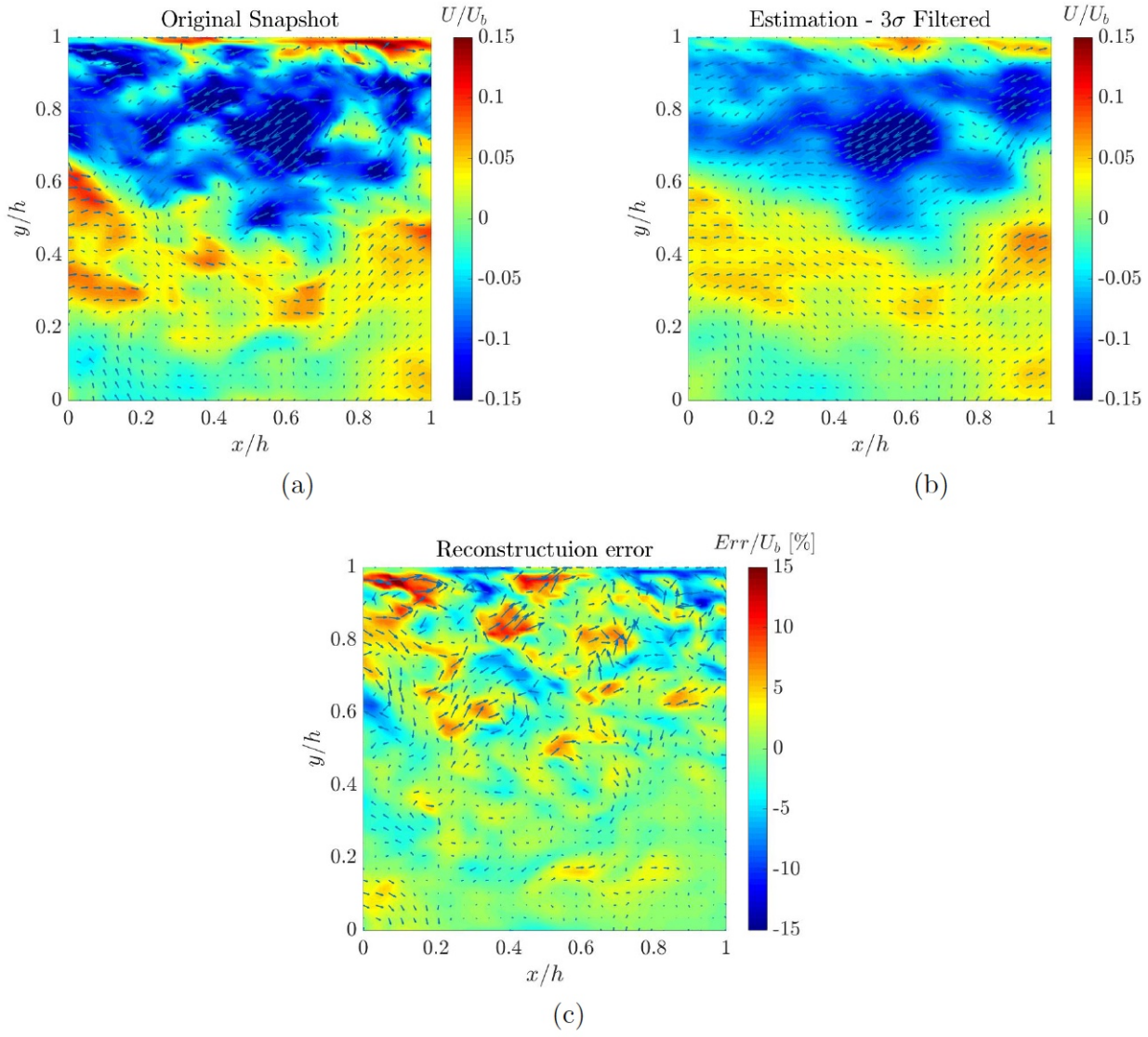

Figure 4.3: Original snapshot, dynamic estimation with case 33 and resulting instantaneous error at an arbitrary time t. The 3σ filtered. a) Original. b) Estimation. c) Instantaneous error.

The comparison between the original snapshot obtained at an arbitrary time instant t and its dynamically estimated reconstruction are shown in figure 4.3 above, together with the resulting instantaneous error. Note that it was decided to use adimensional quantities to depict the different velocity fields and for that, the channel bulk velocity U_b was used. This rule applies to from now on. The field estimation allows the observation of high scale structures in the flow, such as the vortical features in this example. The areas with the highest reconstruction uncertainties, reaching peaks of $\pm 15\%$ of the bulk velocity, concentrate in the regions closer to the wall, where the smallest scales of the flow are poorly estimated. This can be due to the difference in the convective time related to the interaction of the flow with the

wall and that impedes the correlation between probe and field signals. Also notice that the reconstruction performed here concerns a 2D slice and thus any structure affected by the three dimensional character of turbulence may be convected to the out-of-plane regions. Notwithstanding the above, it is seen that the dynamic estimation is able to accomplish the ultimate goal of obtaining a fair description of the largest coherent structures in the flow.

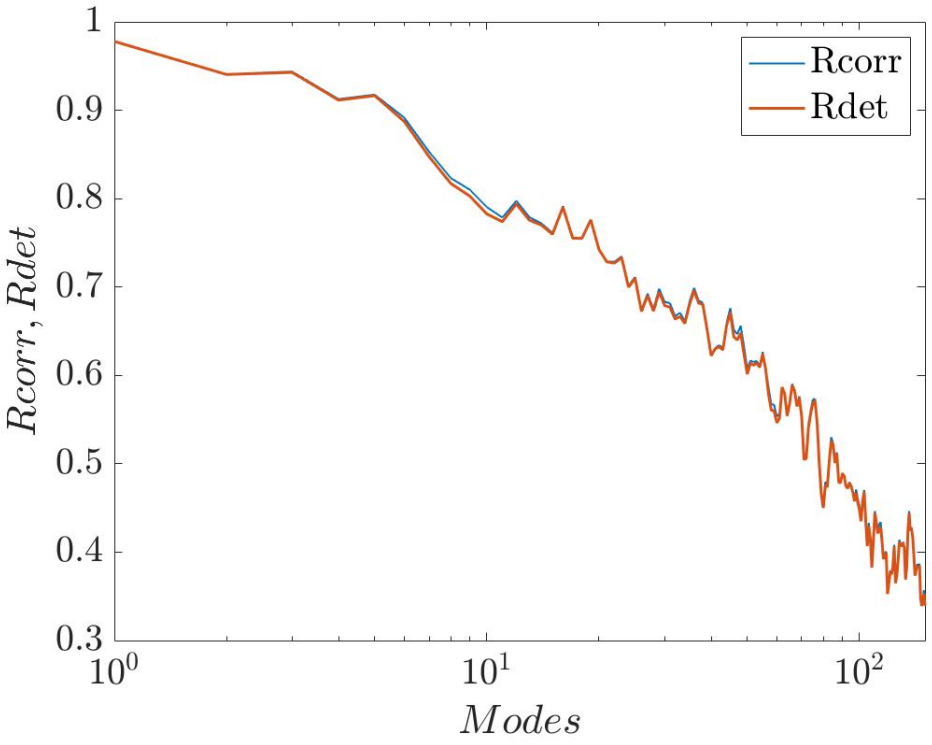

Figure 4.4: Correlation and determination coefficients. The 3σ rule was used for data filtering.

The estimation accuracy between the estimated time coefficients and the real ones is assessed by means of two parameters, the correlation and determination coefficients defined in section 4.1. By looking at figure 4.4 it can be seen that despite the more conservative character of the determination coefficient it shows a great resemblance with the correlation coefficient thanks to the application of the filtration criteria. In line with the above results better estimation results were acquired for the more energetic modes, throwing correlation values higher than 0.5 for the first hundred of modes. On the contrary, the correlation of smaller motion scales falls down to values between 0.35-0.45 which justifies the larger estimation error showed before for these features.

For the completeness of the parametric study performed upon case 33, the average estimation error (equation 4.6) and the average velocity correlation (equation 4.7)

for the same time instant t above were obtained. These contours, shown in figure 4.5 demonstrate the coherence of the dynamic estimation results, being complementary to each other. The snapshot points with the largest mean estimation errors match those showing the poorest average velocity correlation and vice versa. Also, the correlation between probe and field measurements is found to be more intense in the proximity of the virtual probes, growing weaker as the distance between the estimation point and the probes increases. Also the distorting effects of the channel wall are once more reflected in terms of a peaking average error in the nearby area.

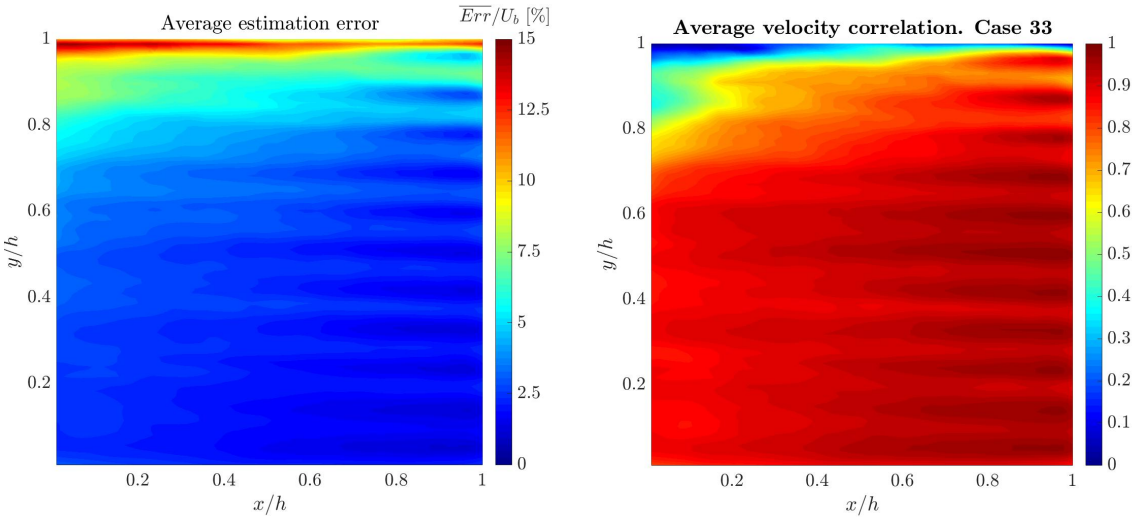

Figure 4.5: Average estimation error and average velocity correlation for case 33. 3σ rule filtered. a) Average estimation error. b) Average velocity correlation.

4.3 Number of components

Figure 4.6 shows the correlation coefficient of the estimated and real temporal coefficients for the six different combinations of field and probe measurements considered in this study. According to such results, the reader can verify the choice for the most robust case made in the previous section. This way, the case 33 is the one providing the strongest correlations, especially from mode 15 onwards. Cases 22 and 32 also provide high correlation values while a significant correlation weakening is observed in the results obtained with combinations 12,13 and 23. Such difference remarks the importance of the amount of velocity components that the virtual probe is able to provide. The three cases involve the establishment of the correlation between a field signal that provides the planar or even a 3D description of the velocity at each point with virtual probes that acquire velocity measurements with at least one dimension less. This correlation dilution is even more severe in the cases where the field signals distant 2 dimensions from the ones procured by the virtual probes.

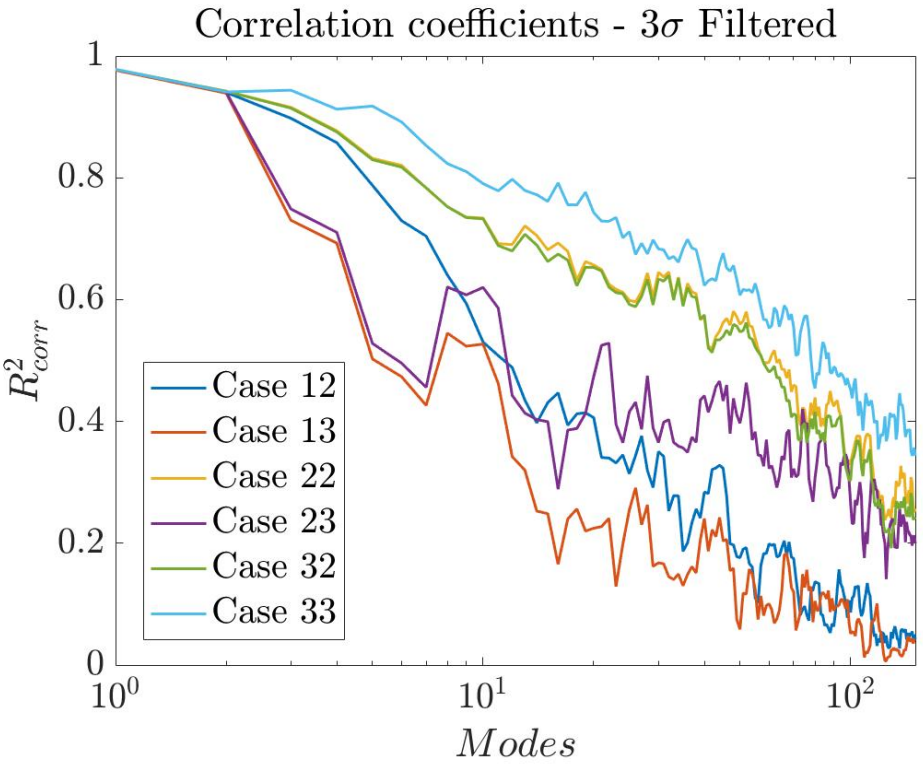

Figure 4.6: Correlation coefficient obtained for different combinations of field and single-point measurements.

To support these results the contours depicting both the average estimation error and average velocity correlation for each of the previous cases are included in the following (see figures 4.7 and 4.8). It is easier observed how the average estimation error increases as the number of velocity components acquired by the virtual probes becomes smaller. Again the average correlation contours show a complementary behavior with respect to the obtained average error.

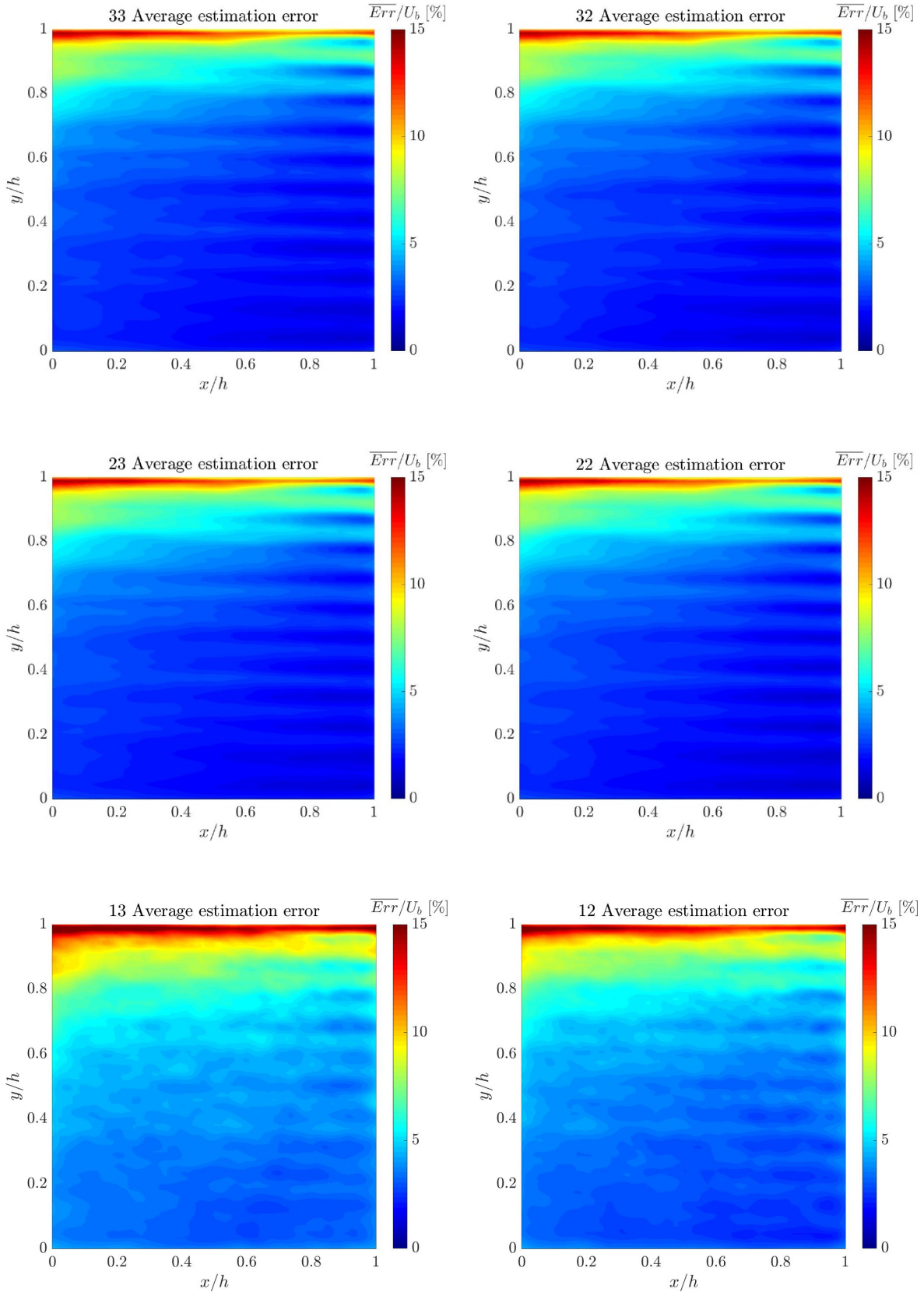

Figure 4.7: Error estimation for all combinations of velocity components.

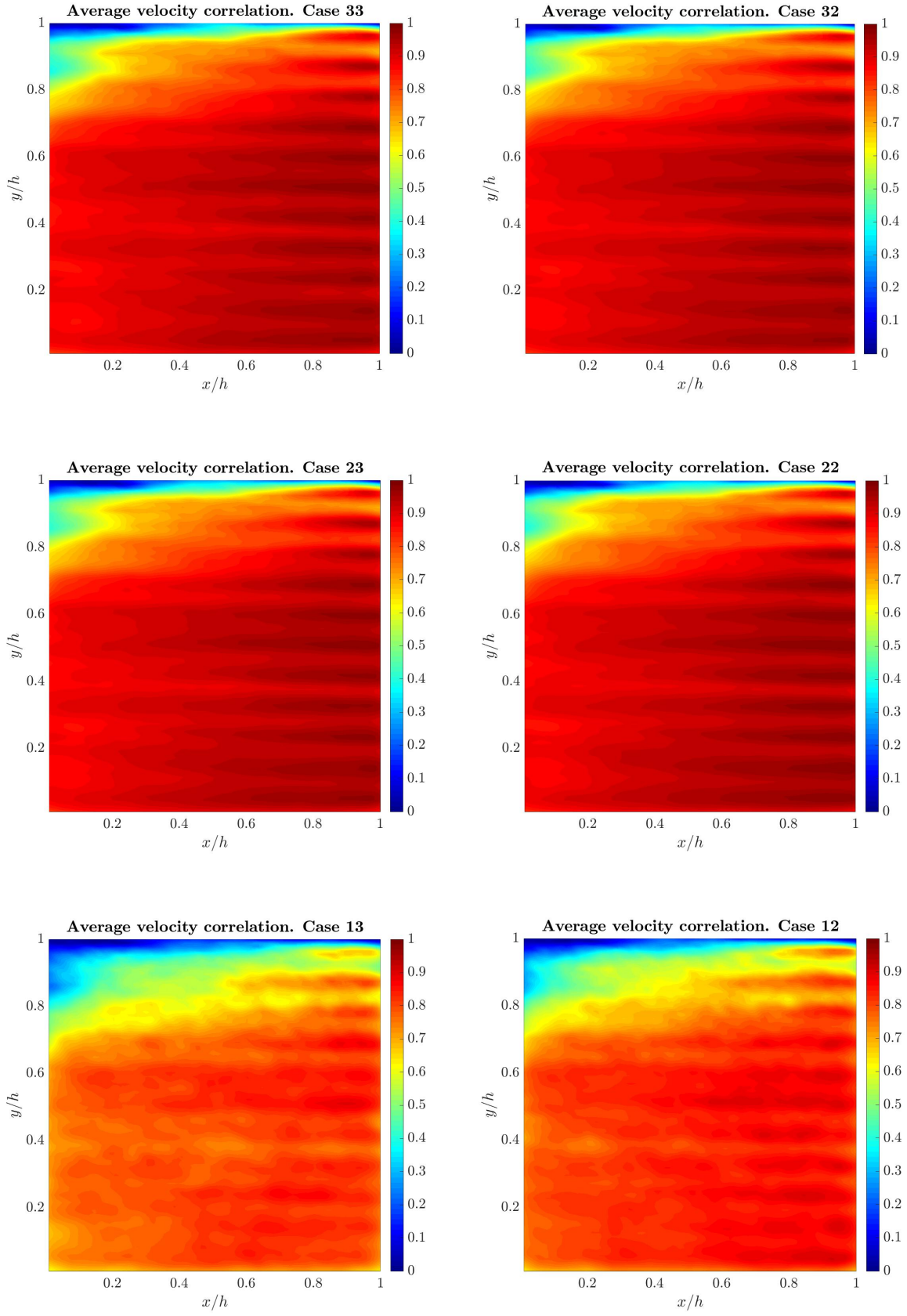

Figure 4.8: Average velocity correlation for all combinations of velocity components.

For simplicity of comparison, the same arbitrary time instant t reconstructed in the previous section for the robust case 33 was also reconstructed with the case 12. The result is shown in figure 4.10 while figure 4.9 shows the original snapshot. The reason behind is that it represents the most common real measurement scenario, that is the availability of 2D field snapshots to be combined with probes acquiring just one velocity component of the flow.

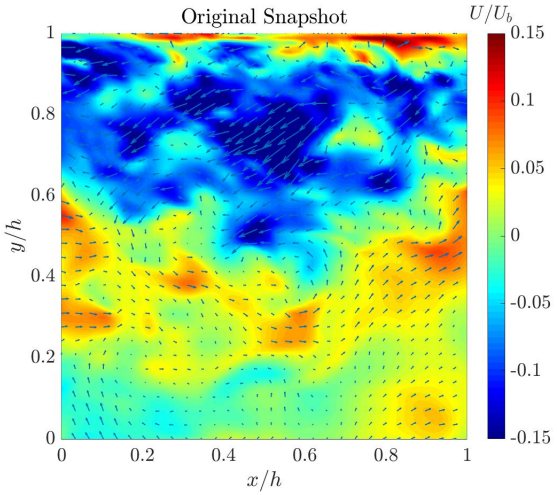

Figure 4.9: Ground truth velocity snapshot at a time t.

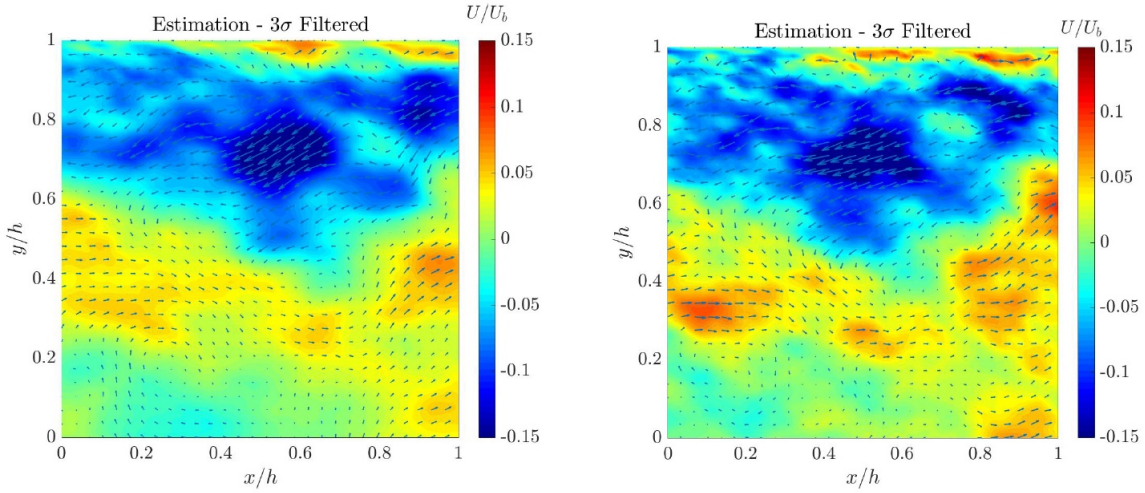

Figure 4.10: Snapshot reconstruction. The 3σ rule was used for data filtering. Left: Case 33. Right: Case 12.

Contrasting the estimations performed with 33 combination and the 12 case depicted above it can be seen a general deterioration of about $\pm 5\%$ in the estimation of the

flow field, what matches the behavior predicted from the tendency of the correlation coefficients in figure 4.6. The reconstruction errors for both cases are collected in figure 4.11 The robustness of the combination of three-dimensional field and probe measurements is reaffirmed while the peaking uncertainties condensate again around the wall in both cases as mentioned before.

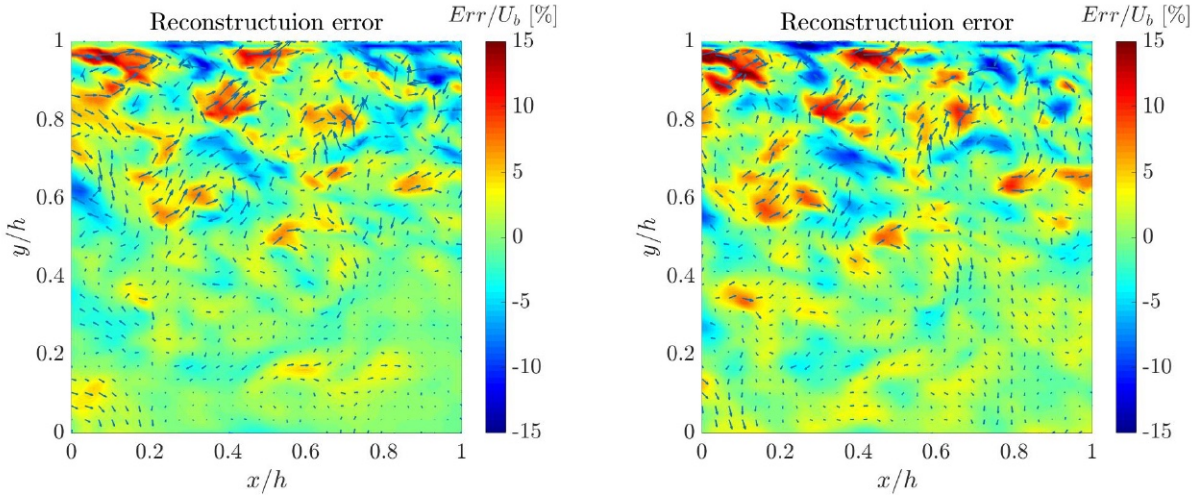

Figure 4.11: Instantaneous error reconstruction. The 3σ rule was used for data filtering. Left: Case 33. Right: Case 12.

4.4 Noise effects

The presence of noise contamination upon the dynamic estimation of the velocity field is analyzed here. Random noise of different intensities was added to the virtual probe and field snapshots and first no truncation criterion is applied in order to isolate the impact of noise presence. Later, the standard truncation criterion 3σ is imposed for a further analysis. Again the case 33 is chosen for the parametric study.

4.4.1 Noise effects in the absence of data filtering

Several noise levels were imposed to the synthetic data. In particular, a 3,5 and 7% of the field turbulent fluctuations were added to the reference field and probe data used for the computation of POD modes and correlation matrices, as well as to the single-point measurements on which the dynamic estimation of the out-of-sample snapshots was based.

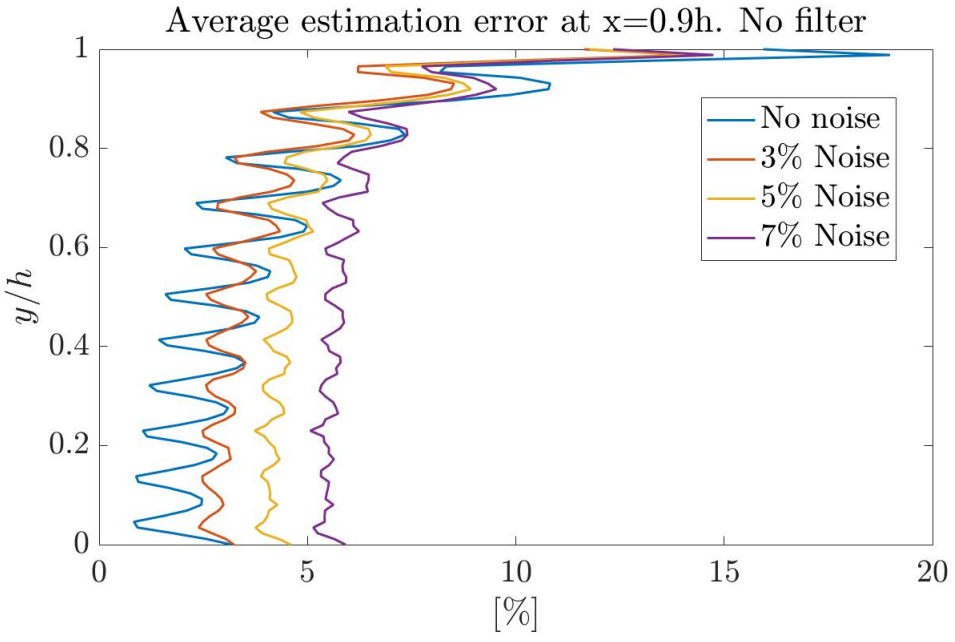

Figure 4.12: Average estimation error at x=0.9h. Non filtered data.

The average estimation error is chosen as the figure of merit for the illustration of the noise upon the performance of the dynamic estimation. Figure 4.12 shows the resulting average error profile at the coordinate x=0.9h for the cases of clean data and the consideration of 3 different noise levels in ascending intensity. The choice of case 33 without applying any truncation rule allows to isolate the detrimental effects of noise upon the already highlighted robustness of this case. As expected,

the introduction of noise reduces the accuracy of the estimation as the contamination intensity increases.

The wavy shape of the profile is due to the higher correlation found at the location of the 11 probes and the immediately surroundings. This tendency is not lost as one gets closer to the channel wall but a significant increase in the uncertainty is observed and the inaccuracies throw closer values to each other. The reason behind is the existence of smaller motion scales resulting from the interaction of the turbulent flow with the solid boundary. This effect has a stronger influence on the estimation performance at that region that the proper location of the probes but it adds to the wavy trend to conform the profile shape displayed at the top part of the figure above.

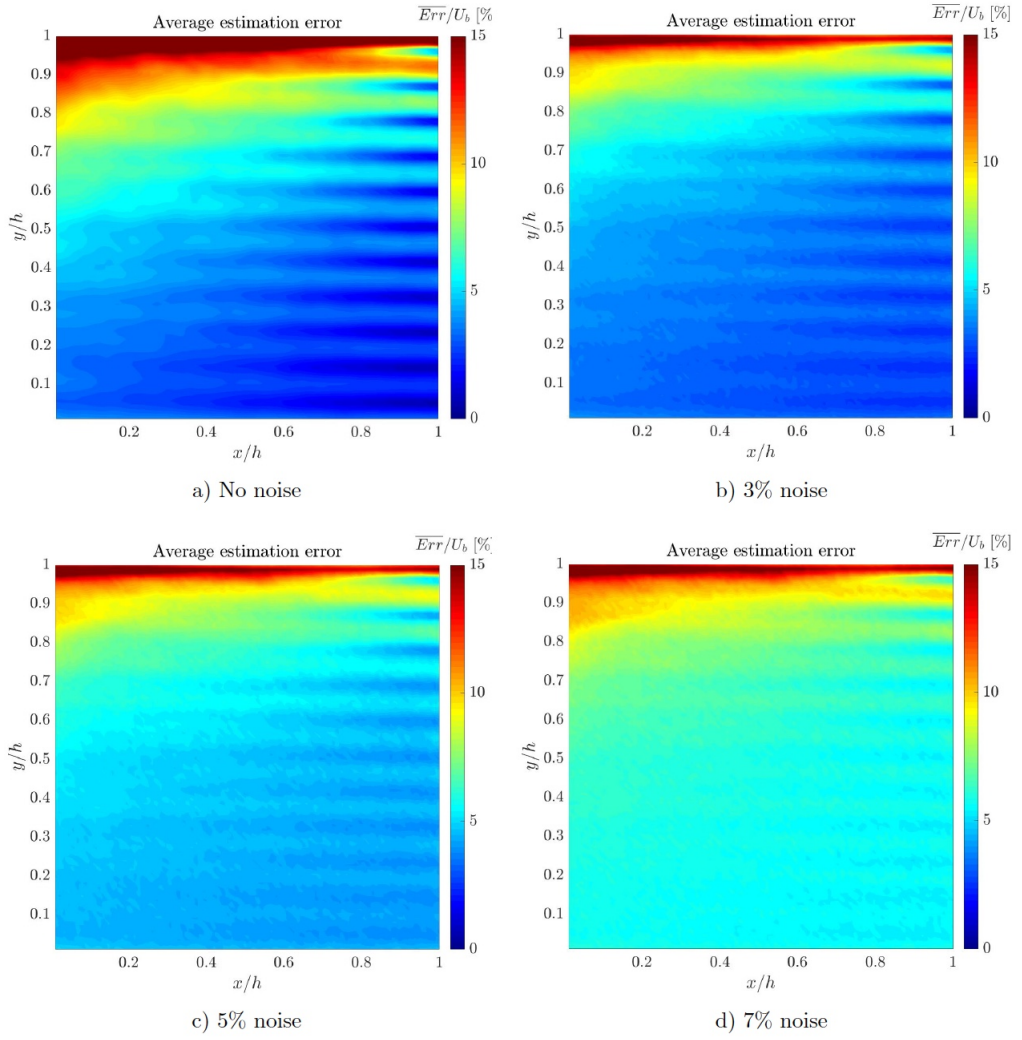

Figure 4.13: Average estimation error for several noise levels. Case 33 with no data filtering. a) No noise. b) 3% noise. c) 5% noise. d) 7% noise.

A global view of the already described behavior is offered in figure 4.13. For a low intensity noise as the case of 3% the estimation still provide acceptable performance values in the regions close to the probes and far from the wall. However, for a 5% and 7% of noise presence the average estimation error increases notoriously, making clear the need for a method that enables the removal of this undesired contamination. Notice also the aforementioned wavy velocity contour that arises close to the probes locations.

4.4.2 Noise effects on filtered data

Once the need for a filtration criterion has been justified, here the effects of noise were introduced in order to study their net impact upon filtered data. The following figure shows the correlation coefficient for case 33 filtered with the standard 3σ rule.

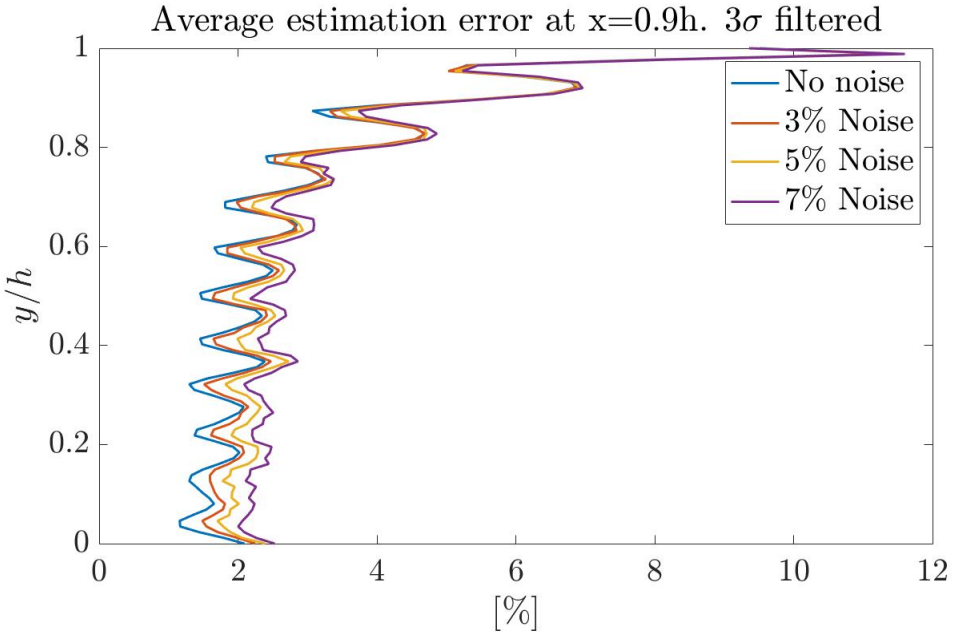

Figure 4.14: Average estimation error at x=0.9h. All cases were filtered with the 3σ rule.

In figure 4.14 he filtrating effect of the standard truncation criterion stands out. The elimination of the major part of uncorrelated data results in the reduction of the average error and the compaction of the curves while the dominant near-probe correlations undulating profile shape is recurrent. This effect is related to the high impact of the truncation in comparison to the low noise intensities used to disturb the estimation and which in fact very similar to each other. The truncation shows poor effectiveness at the wall surroundings since the uncertainties arising there are not driven by noise effects, as explained previously.

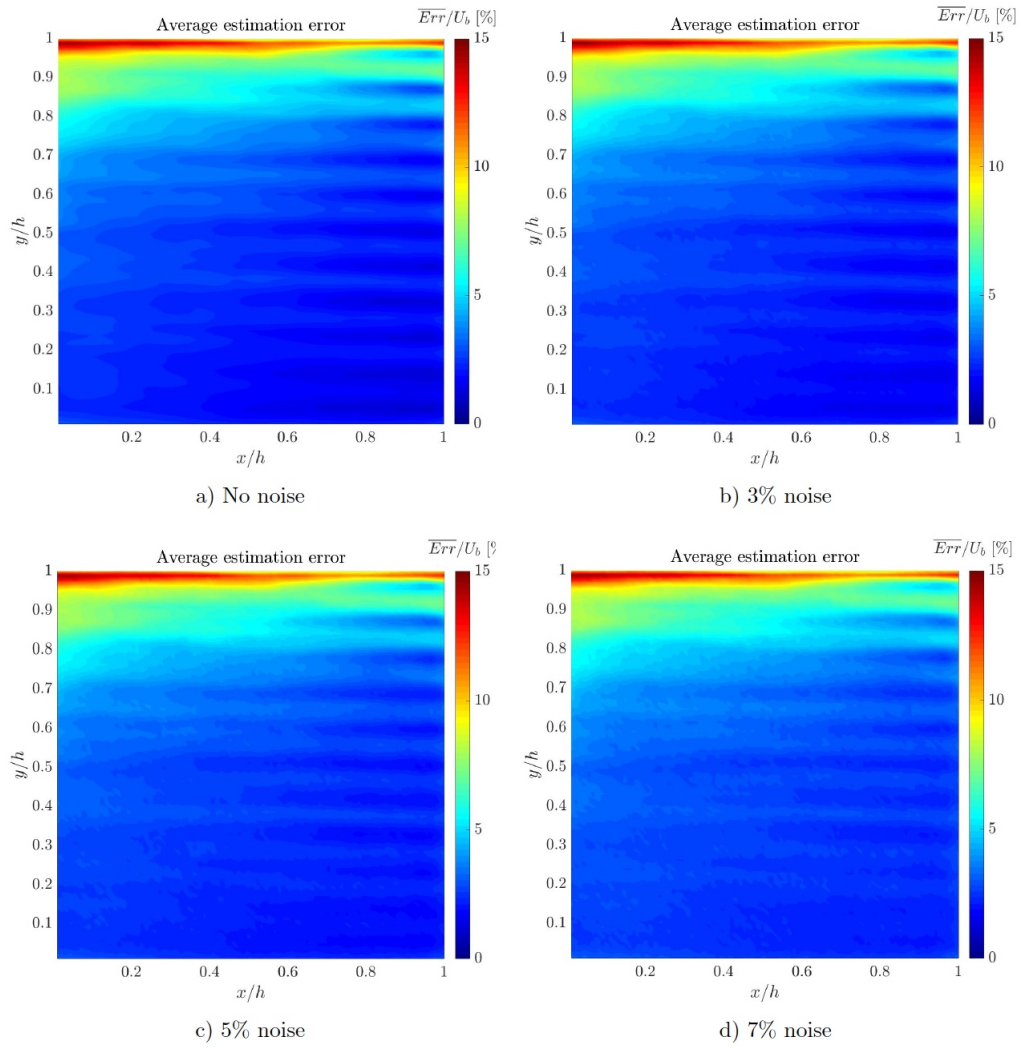

Figure 4.15: Average estimation error for different level noises. Case 33 filtered with the 3σ rule. a) No noise. b) 3% noise. c) 5% noise. d) 7% noise.

The highlighted improvements achieved after the implementation of the 3σ truncation criterion are more clearly demonstrated in the complete contours of the average estimation error in figure 4.15. No significant differences are found between each of the noise-contaminated cases and the clean data instance. The truncation impact is more powerful than any of the noise intensities considered, which again proves the need for its implementation in order to ensure higher estimation accuracies.

4.5 Truncation criterion

The final part of this parametric study addresses the validation and optimization of a truncation criterion for the removal of the non-correlated part of the signals. To that end, the so far adopted 3σ rule used for data filtration was evaluated and, departing from this standard, two other alternatives, one more conservative (4σ filtration) and another one less restrictive (2σ filtration) were also tested.

4.5.1 Truncation assessment for clean data

In the absence of noise, the correlation coefficients computed for the standard filtration and the two variants proposed here are depicted in figure 4.16 based on case 33. The benefits associated to the removal of the uncorrelated signals is obvious. Higher levels of temporal correlations are achieved by means of even the weakest of the truncation criteria tested.

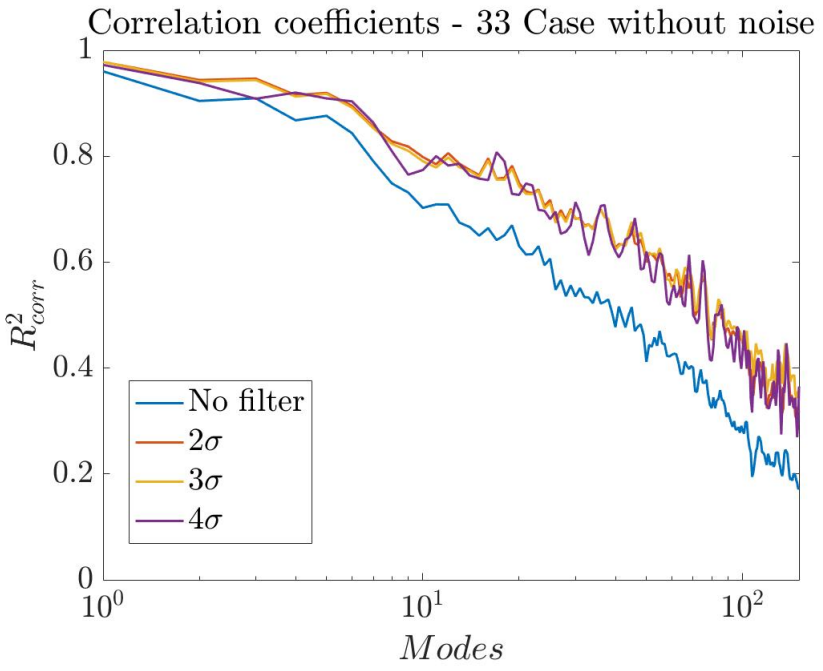

Figure 4.16: Case 33. Correlation coefficient for different truncation criteria in the absence of noise effects.

One could deduce that the more aggressively filtered the data is, the more accuracy could be achieved. However the previous picture shows that this is not necessarily true since already with the use of the 3σ rule the vast majority of the non-correlated signal can be eliminated ($\approx 99.7\%$). It can be also seen that even the removal of correlations in the range of $\pm 2\sigma$, which represents a weaker truncation of the signal

with respect to the standard 3σ rule, high performances can be achieved. In fact, going a step further towards the imposition of a more demanding criterion is not found to be efficient since just about 3% of the noise is likely to be found in the range $-4\sigma \leq \Xi_{ij} \leq -3\sigma$ and $+3\sigma \leq \Xi_{ij} \leq +4\sigma$, as shown by the assumption made in section 2.3.2.

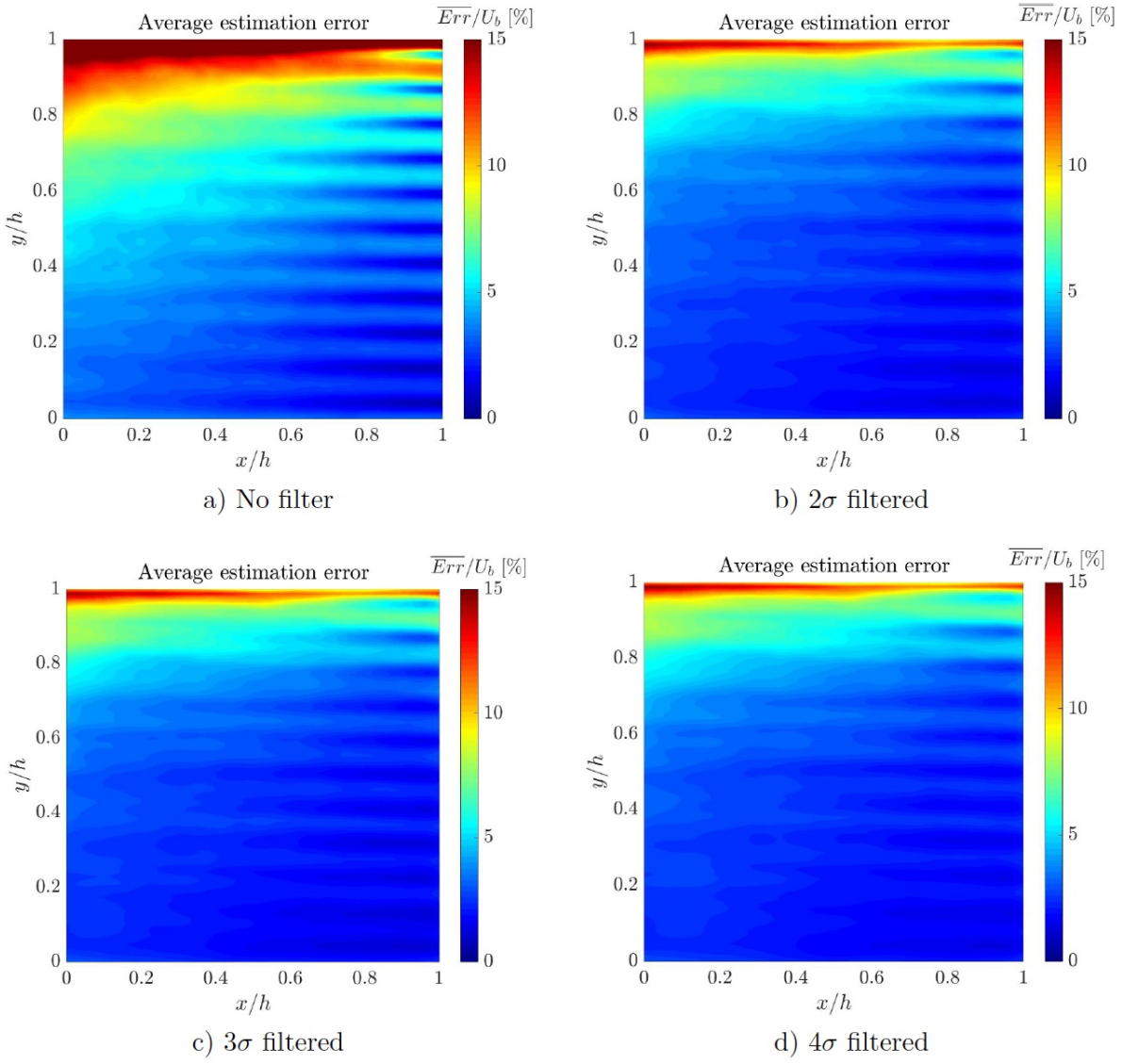

Figure 4.17: Case 33. Average estimation error for different truncation criteria in the absence of noise. a) No filter. b) 2σ filtered. c) 3σ filtered. d) 4σ filtered.

The mapping of the average estimation error for the different filtering intensities in figure 4.17 corroborates the results previously illustrated by the correlation coefficient. The critical region continues to be the area close to the channel wall, where the dynamics of the flow are disturbed because of the interaction of smaller motion

scales in the free moving fluid with the solid boundary. Even when account is taken of this, the truncation of the correlated signals considerably improves the estimation nearby that area, being especially noticeable with the use of 2σ and 3σ filters. It is also remarked that no significant benefits are extracted from the hardening of the truncation threshold up to 4σ.

Taking again the average error profile at x=0.9h (see figure 4.18), these results are even more obvious and the error reduction is considerably once any of the truncation criteria is used. Since no noise effects were introduced in this case, the performance of all the truncation criteria proposed shows to be very similar and just distinguishable from the non filtered estimation.

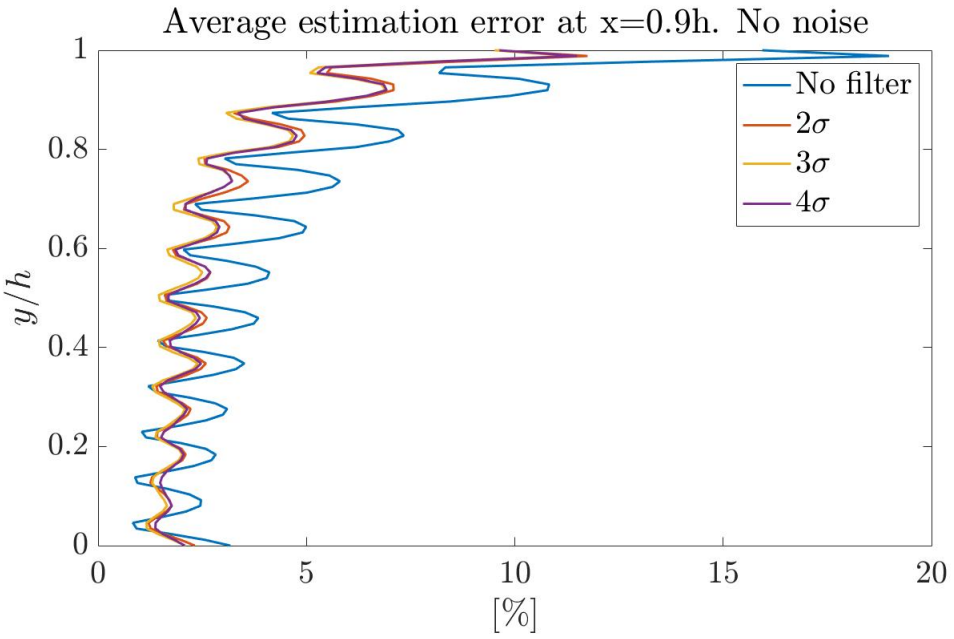

Figure 4.18: Average estimation error at x=0.9h for different truncation criteria.

4.5.2 Truncation assessment under noise effects

The benefits provided by the implementation of a truncation criterion have been confirmed. The next step is to test the performance of the different truncation thresholds proposed under the presence of the noise contamination which is commonly faced when conducting experimental velocity measurements. First, the correlation coefficient is used to assess the correlation between the real time coefficients and the ones dynamically estimated. The result can be seen in figure 4.19 below. As usual, the test case is chosen to be the robust 33 which was tested under the most demanding noise conditions here considered, i.e. with an addition of 7% of the turbulence intensity as noise.

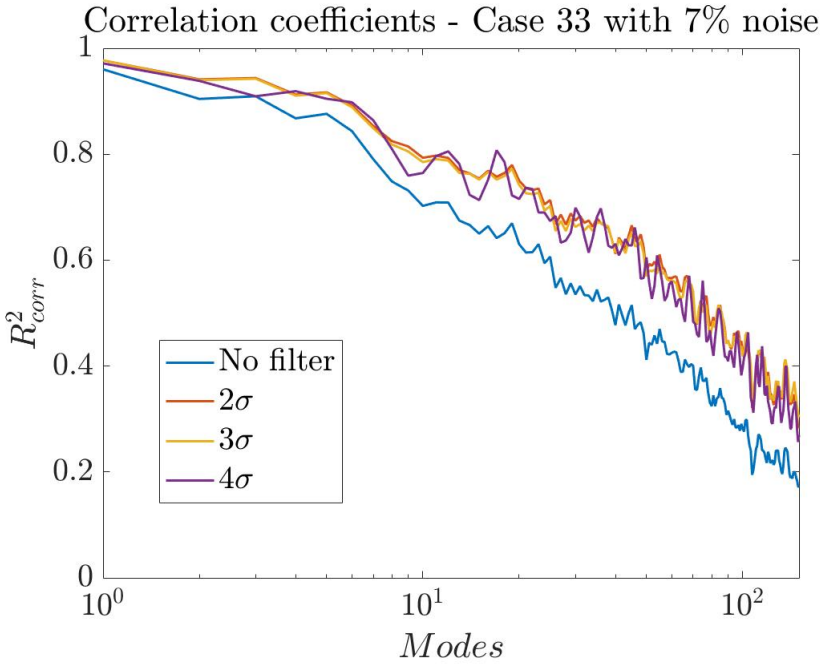

Figure 4.19: Case 33. Correlation coefficient for different truncation criteria. 7% noise addition.

Despite the enforcement of the highest noise contamination initially considered in this study, the robustness of case 33 is able to compensate its spurious effects and no remarkable effect is observed, especially for the first POD modes which accumulate most of the turbulent energy of the flow. On the other hand, the truncation criterion does not seem to loose its refinement potential, providing similar results to the corresponding cases without noise addition considered before. Thus, the noise effects upon the performance of the filtering rule can only be poorly observed, making again a difference just with respect to the non filtered data.

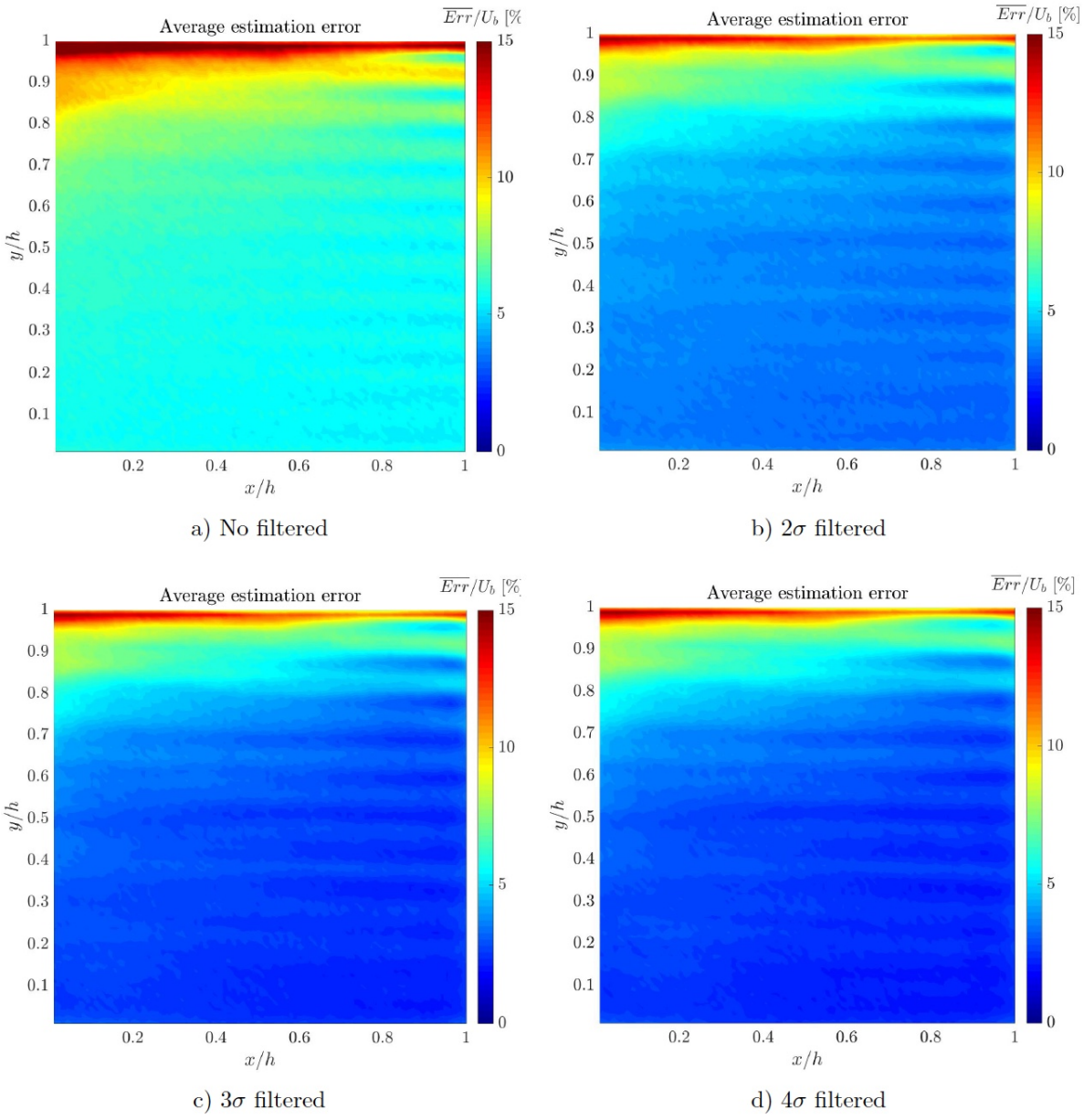

Figure 4.20: Case 33 with a 7% of noise addition. Average estimation error for different truncation criteria. a) No filter. b) 2σ filtered. c) 3σ filtered. d) 4σ filtered.

For the sake of completeness, the average estimation error maps with the imposition of noise (figure 4.20) and the usual average error profile (4.21)close to the probes spanwise location (x=0.9h) are included here. The truncation shows the same behavior described so far, the use of the truncation rules effectively improve the estimation of the fluid field but provide no important performance benchmarks compared to each other.

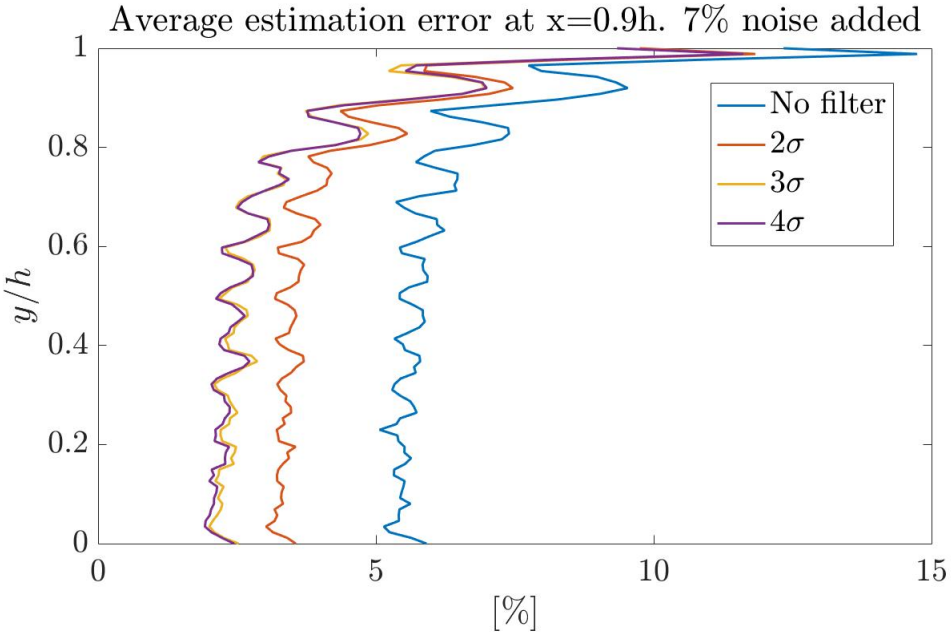

Figure 4.21: Average estimation error at x=0.9h. No noise.

4.6 Higher noise intensities

According to the results discussed so far, the standard 3σ rule to be the most robust and efficient filtration threshold. In the following, the reader can find the results obtained using case 33 filtered with such criterion and under high noise intensities ranging from 11-25% of the velocity fluctuations.

Figures 4.22 show the correlation coefficients obtained for the four high noise levels simulated. It must be noticed that the imposition of such extreme noise intensities compared to the turbulent fluctuations present in the flow do no represent a real case scenario and has been used here just to prove the limits of the proposed filtration criterion.

The large intensities of noise imposed to the velocity measurements do not disturb in the robust character of the standard truncation, which continues to be clearly dominant for the first 100 modes. From the 100th mode on, the strong noise contamination becomes more evident as the motion scales become smaller, i.e. as the mode number increases. However, this decrease in correlation found in the higher modes is not found to be very sharp compared to the extreme noise intensities that they represent. This is be due to the fact the dynamic estimation performed here relies in the most optimum conditions possible: the use of a database with a high quality regarding statistical convergence (large amount of modes) and the use of the

most robust case (33). In a real case scenario noise intensities are not usually so large but at the same time the statistical quality of the measurement data is lower, or the number of velocity components acquired for the estimation is different from the 33 combination used here. Thus, in real conditions the quality of the estimation is lower.

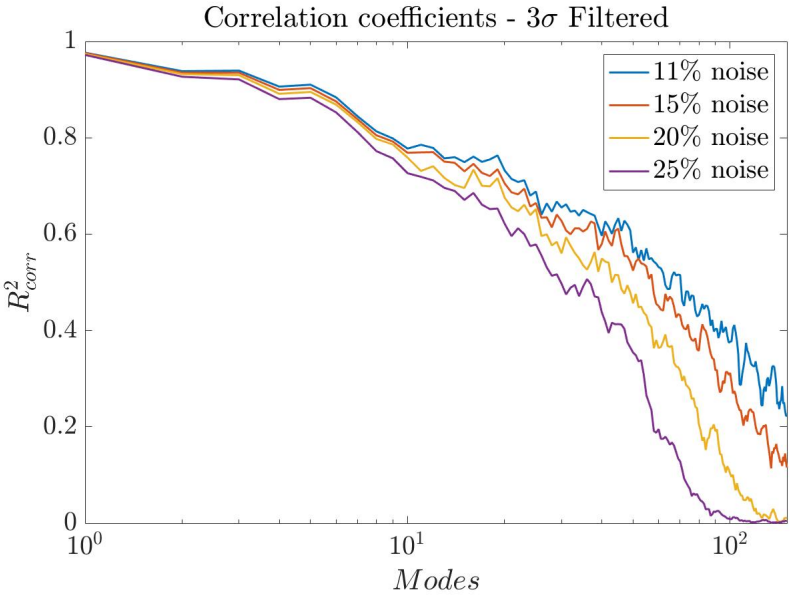

Figure 4.22: Correlation coefficient. Case 33 filtered with 3σ rule.

The same behavior explained before is observed in figure 4.23, which depicts the average estimation error profile at the usual location x=0.9h. The maximum error is found again in the region close to the wall and its module is not affected by the large noise intensities imposed since they are related to different phenomena. The wavy shape of the error profile associated with the location of the probes is notoriously disturbed by the noise effects but still the quality of the results is very good in comparison with the intense noise presence (3-5% average error).

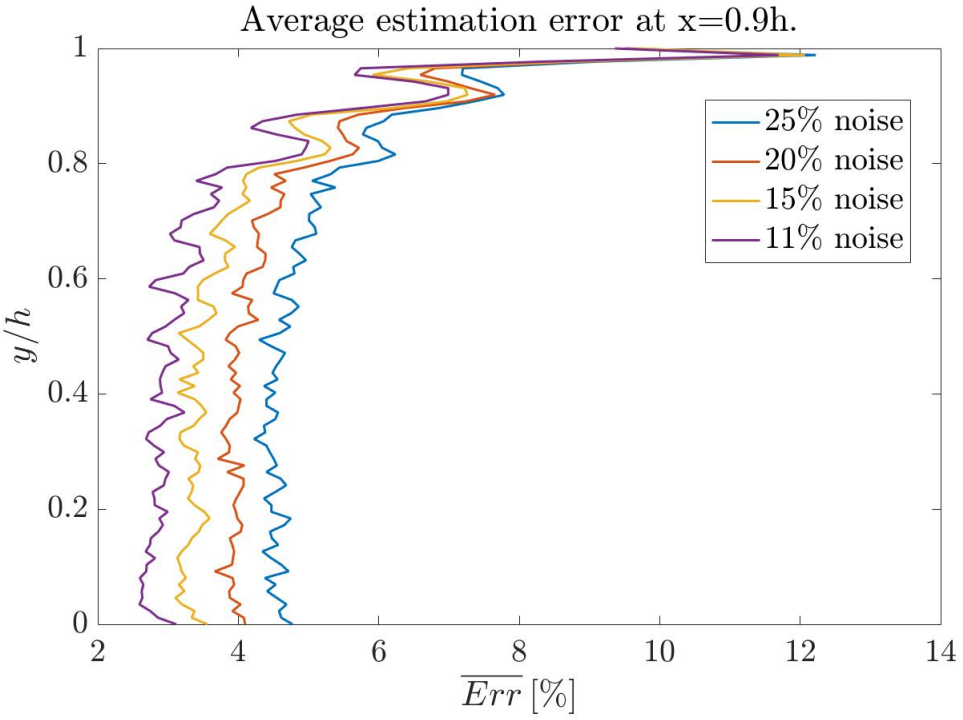

Figure 4.23: Average estimation error profile at x=0.9h. Case 33 filtered using 3σ rule.

To finish, the average estimation error contours for the different noise intensities are collected in figure 4.23. These error maps corroborate the behavior predicted by previous figures and reaffirm the robustness of case 33 with the use of the standard 3σ filtering rule.

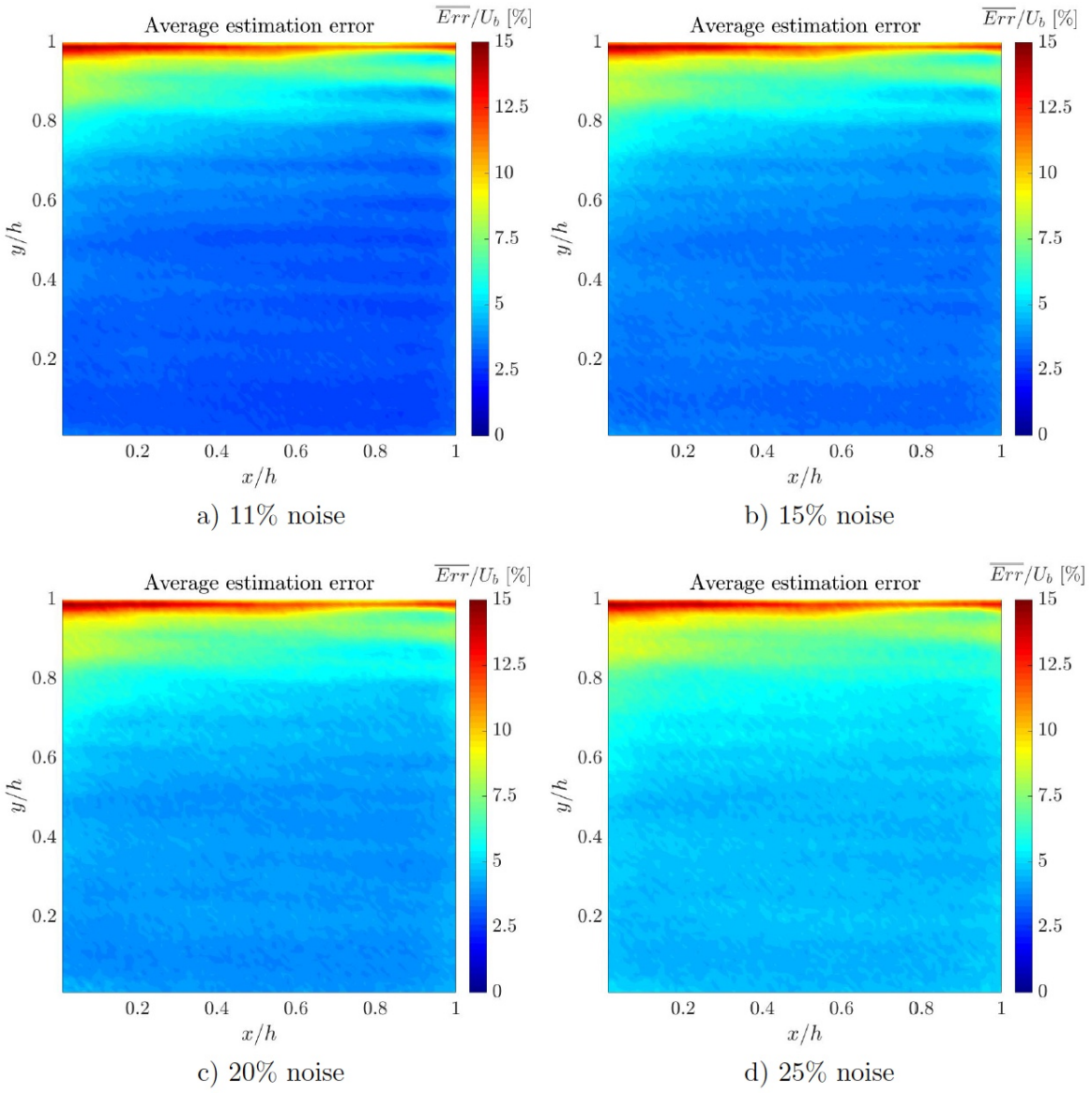

Figure 4.24: Average estimation error for case 33 filtered with 3σ rule.

Chapter 5

Conclusion

The present work focused on the study of turbulence through the assessment of a method for the estimation of time-resolved turbulent velocity fields starting from the combination of non-time resolved velocity fields and time-resolved point-measurements. Specifically, the aim was to validate and refine this technique with establishment and evaluation of a truncation criterion for the removal of uncorrelated signals, with the aim of allowing a clearer and simpler identification of the turbulent coherent structures in the flow. Continuing with the work developed by Discetti et al.[1], the dynamic estimation of the velocity field using Extended Proper Orthogonal Decomposition was implemented and validated. On the other hand, a parametric study was developed in order to evaluate the already proposed standard filtration threshold under the presence of noise and for several field-probe velocity components combinations.

The approach validation was performed using a Direct Numerical Simulation of the flow through a rectangular channel expecting to show high spectral richness. Several truncation criteria besides the already proposed standard 3σ rule were implemented. All those truncation thresholds are supported by the assumption of a Gaussian distribution of noise and uncorrelated signals in the correlation matrix built from the field and probe POD's time coefficients.

The results obtained lead to several conclusions:

- The relation of velocity components between the field and single-point measurements used for the obtaining of the correlation matrix plays a role the performance of the dynamic estimation based on EPOD modes. Defining a velocity component ratio

$$r = \frac{\#\ velocity\ components\ measured\ by\ the\ single\ point\ probes}{\#\ velocity\ components\ measured\ by\ the\ field\ samples}, \qquad (5.1)$$

the resulting average error is reduced for estimations performed with ratios r≥1. This way, the most robust cases turned out to be 33, followed by cases 32 and 22, while options 23,12,13 showed lower performances.

- The use of a correlation truncation criterion when dealing with turbulent flows with large spectral richness provides substantial accuracy improvements. The standard 3σ rule proved to be the most suitable filtration threshold in terms of efficiency and effectiveness for the removal of uncorrelated signals and mitigation of low intensity (up to 7% of the turbulent fluctuations) noise effects. Despite the initial idea of a possible improvement of the results derived from a more restrictive filtration criterion, the here also tested 4σ truncation rule did not demonstrate remarkable profits but rather seemed to absorb not just the uncorrelated part of the signals but also some valid correlation data, thus overstepping its filtration purposes.

- The great performances shown by the 3σ rule under the presence of large noise intensities are far from representing a real scenario in which the quality of the statistical convergence and the measurement conditions are not always the most optimum.

Notwithstanding the above, the powerful truncation approach validated in this study was only tested under low noise intensity conditions and the question of its useful implementation range for higher intensely contaminated data remains open.

Chapter 6

Project Environment

6.1 Socioeconomic impact and regulatory framework

In principle, no regulatory framework is needed to cover the work done since no experimental test were carried out.

With regard to the socioeconomic impact, it has been repeatedly reminded that turbulent flows do not represent an anomaly but rather a common phenomena in our everyday surroundings. One can describe some turbulent flows from the qualitative point of view in order to depict the concept of turbulence. However, its mathematical description is needed in order to take advantage of the prediction of its dynamic characteristics when desired and also to mitigate its negative effects when required.

This way, the accurate caption of the turbulence dynamics could be used to, for example, improve aerodynamic, propulsive, heat-tranfer, matter mixing and industrial processes in general. It could also provide a means for a better control and mitigation of environmental pollution.

One of the most remarkable potential applications of turbulence prediction and modeling is the flow control in aerodynamic processes. The complete understanding of turbulent motion could be used to predict and thus mitigate the high level of drag associated to such events by means of enhanced aerodynamic profiles, increasing the efficiency in air transportation. Also combustion processes could benefit from the control of turbulent flows, allowing for a better matter mixing and thermal stresses control that could increase the economic performance of jet engines.

These improvements would also impact the aircraft fuel consumption and thus the amount of polluting substances that are expelled to the atmosphere as a result of the daily activities of this sector. In fact, there are some programs already in progress today which are dedicated to push technological initiatives that may alleviate the adverse effects of air transport contamination towards the design of cleaner aircraft.

One example is the European Joint Undertaking called Clean Sky, which is committed to reduce gas emissions, noise levels and also the production of CO2 as part of the EU Horizon 2020 research and innovation programme. More information about this initiative can be found in the program website *http://www.cleansky.eu/*. Therefore, any research input to the better comprehension of turbulence events can contribute to a more efficient and eco-friendly aerospace industry.

6.2 Project cost

The purpose of this section, rather than pretending to expose an exhaustive computation of the project budget, is to perform a drilldown on the main sources of costs that could arise as a result of the development of this or other similar projects. Below is a summary on the core expenses ascribed to this work:

- ***Software.*** In this particular case the software environment employed for the development of the algorithm was Matlab. The average price for the obtaining of a Matlab academic type license with perpetual character is 500€. However, there is also the possibility of using similar free-software resources such as Python or GNU Octave. On the other hand, the most common operative system is Windows and one of its latest, high-performing versions, Windows 10 Pro, has a price of 259€.
- ***Labor force.*** The amount of time dedicated to the present work is roughly estimated to be 360h. The minimum cost associated should be estimated relying on the legally stipulated base salary for a Graduate Aerospace Engineer which depends on country regulations. In this case, for an estimated average cost of 20€/h the total labor costs amount 7200€.
- ***Hardware.*** The only hardware resource used here was a laptop and the resulting costs is linked to its depreciation during the project development time. With an approximately initial cost of 850€ and assuming a linear depreciation for an average lifetime of 4 years ($\approx 35.040h$) the depreciation is equal to 0.024€. The resulting cost for the 360h of work amounts a total of 8.70€.
- ***Various costs.*** Other possible arising costs are office materials, electricity or Internet connection were roughly estimated to be 400€.

A summary of the costs described so far is collected in table 6.1.

TOTAL PROJECT COST	
Concept	**Cost**
Software	759€
Labor	7200€
Hardware	8.70€
Various	400€
TOTAL	8367.7€

Table 6.1: Total project costs

6.3 Project timeline

The development of the project has spanned over almost 10 months. During this time period the kind of dedication to it was mostly part-time and was progressively increased proportional to the project stage complexity. The acquisition of background knowledge about the problem and literature general review took approximately 50 hours of reading spread into two months with some brief dead-times in between due to holidays and final examinations season.

The second stage of this project was the development of the core algorithm for the dynamic estimation. A total of 150 hours over 3 and a half months were spent for the code planning, implementation, optimization and also for the running of a small test case used for algorithm validation. Note that this latter task increased the workload since it overlapped with the introduction of the true test case and the design of the parametric study to be performed upon it, as it can be appreciated in figure 6.1.

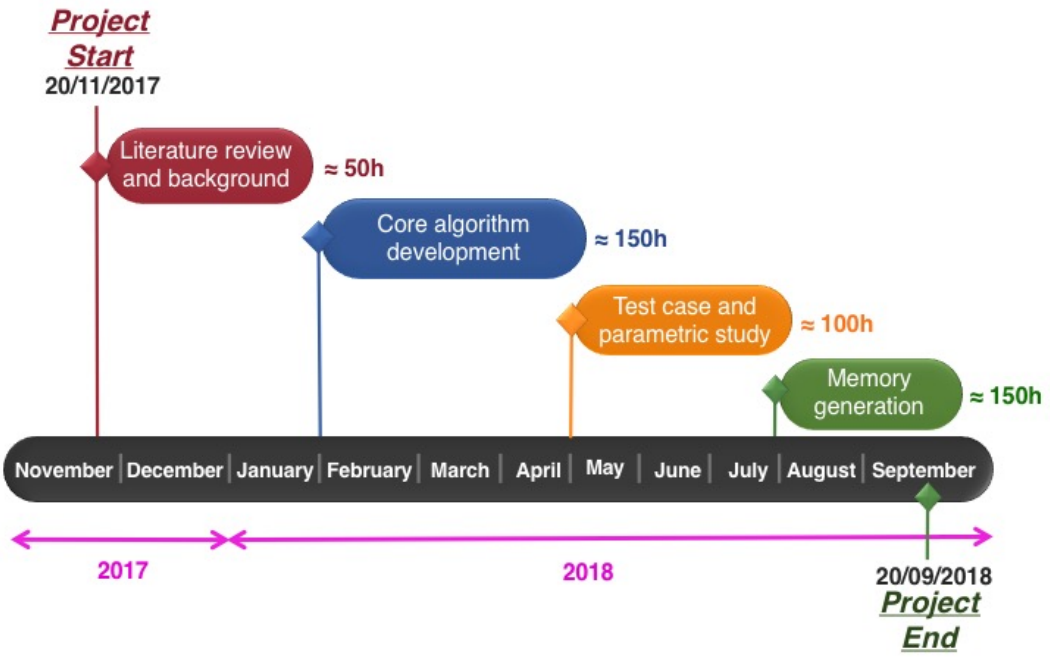

Figure 6.1: Project Gannt chart.

After the proper revision and validation of the Matlab code, the work development continued with the application of the algorithm to the test case in consideration. A complete parametric study of the case was performed by means of the numerical simulation of multiple case scenarios, varying and combining different working conditions for the subsequent analysis of the results. The duration of this phase was estimated to be close to 100 hours with again the appearance of a small time period during which the time dedicated was reduced due to the increasing academic workload between May-June.

The very last stage embraced the study of the results and the generation of the pertinent documentation, which took approximately 150 hours during the last one and a half month of the project schedule. It can be clearly seen how the workload was progressively increased from stage one to the final documentation generation. The project Gannt chart attached above provides a general overview about the project timeline.

Bibliography

[1] Stefano Discetti, Marco Raiola, and Andrea Ianiro. "Estimation of time-resolved turbulent fields through correlation of non-time-resolved field measurements and time-resolved point measurements". In: 93 (2018), pp. 119–130. ISSN: 0894-1777.

[2] Stephen B Pope. *Turbulent flows*. eng. Cambridge: : Cambridge University Press, 2000. ISBN: 0521598869.

[3] *Springer Handbook of Experimental Fluid Mechanics*. eng. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. ISBN: 3540251413.

[4] National Aeronautics and Space Administration. *Navier-Stokes Equations*. NASA. URL: `https://www.grc.nasa.gov/www/k-12/airplane/nseqs.html` (visited on 08/21/2018).

[5] F. Durst. *Fluid Mechanics Developments and Advancements in the 20th Century*. Institute of Fluid Mechanics, University of Erlangen-Nurenberg.

[6] National Aeronautics and Space Administration. *CFD Image of Hyper-X Research Vehicle at Mach 7 with Engine Operating*. NASA. URL: `https://www.dfrc.nasa.gov/Gallery/Photo/X-43A/HTML/ED97-43968-1.html` (visited on 08/24/2018).

[7] *An album of fluid motion*. eng. Stanford: : Parabolic Press, 1982. ISBN: 0915760029.

[8] Stefano Discetti and Filippo Coletti. "Volumetric velocimetry for fluid flows". eng. In: *Measurement Science and Technology* 29.4 (2018). ISSN: 0957-0233.

[9] A. (Arkady) Tsinober. *An Informal Conceptual Introduction to Turbulence*. 2nd ed. Fluid Mechanics and Its Applications, 92. 2009. ISBN: 90-481-3174-X.

[10] F. Michel. "Lärm und Resonanzschwingungen im Kraftwerkbetrieb". In: *Infolge periodischer Strömungsforgänge* (1932).

[11] Jonathan Tu et al. "Integration of non-time-resolved PIV and time-resolved velocity point sensors for dynamic estimation of velocity fields". In: *Experiments in Fluids* 54.2 (2013), pp. 1–20. ISSN: 0723-4864.

[12] C Tinney, L Ukeiley, and M Glauser. "Low-dimensional characteristics of a transonic jet. Part 2. Estimate and far-field prediction". eng. In: *Journal of Fluid Mechanics* 615 (2008), pp. 53–92. ISSN: 00221120. URL: http://search.proquest.com/docview/210896750/.

[13] J. Borée. "Extended proper orthogonal decomposition: a tool to analyse correlated events in turbulent flows". eng. In: 35.2 (2003), pp. 188–192. ISSN: 0723-4864.

[14] Y. G. Guezennec. "Stochastic estimation of coherent structures in turbulent boundary layers". In: *Physics of Fluids A: Fluid Dynamics* 1.6 (1989), pp. 1054–1060. ISSN: 0899-8213.

[15] Dan Ewing and Joseph H. Citriniti. "Examination of a LSE/POD complementary technique using single and multi-time information in the axisymmetric shear layer". In: *IUTAM Symposium on Simulation and Identification of Organized Structures in Flows*. Ed. by J. N. Sørensen, E. J. Hopfinger, and N. Aubry. Dordrecht: Springer Netherlands, 1999, pp. 375–384. ISBN: 978-94-011-4601-2.

[16] Adrian Spencer and David Hollis. "Correcting for sub-grid filtering effects in particle image velocimetry data". In: 16.11 (2005). ISSN: 0957-0233.

[17] V. Durgesh and J. Naughton. "Multi-time-delay LSE-POD complementary approach applied to unsteady high-Reynolds-number near wake flow". eng. In: *Experiments in Fluids* 49.3 (2010), pp. 571–583. ISSN: 0723-4864.

[18] J A Taylor and M N Glauser. "Towards Practical Flow Sensing and Control via POD and LSE Based Low-Dimensional Tools". In: *Journal of fluids engineering* 126.3 (2004). ISSN: 0098-2202.

[19] W. J. Baars and C. E. Tinney. "POD-based Spectral Higher-order Stochastic Estimation". In: *Physics of Fluids* 26:055112 (2014), pp. 1–20.

[20] Zahra Hosseini, Robert Martinuzzi, and Bernd Noack. "Sensor-based estimation of the velocity in the wake of a low-aspect-ratio pyramid". eng. In: *Experiments in Fluids* 56.1 (2015), pp. 1–16. ISSN: 0723-4864.

[21] G Berkooz, P Holmes, and J L Lumley. "The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows". In: *Annual Review of Fluid Mechanics* 25.1 (1993), pp. 539–575. DOI: 10.1146/annurev.fl.25.010193.002543. URL: https://doi.org/10.1146/annurev.fl.25.010193.002543.

[22] V. Jaunet, E. Collin, and J. Delville. "POD-Galerkin advection model for convective flow: application to a flapping rectangular supersonic jet". In: *Experiments in Fluids* 57.5 (2016), pp. 1–13. ISSN: 0723-4864.

[23] L. Sirovich. "Turbulence and the dynamics of coherent structures. I - Coherent structures. II - Symmetries and transformations. III - Dynamics and scaling". In: *Quarterly of Applied Mathematics* 45 (Oct. 1987), pp. 561–571. URL: http://adsabs.harvard.edu/abs/1987QApMa..45..561S.

[24] Jerry Westerweel, Gerrit E. Elsinga, and Ronald J. Adrian. "Particle Image Velocimetry for Complex and Turbulent Flows". eng. In: *Annual Review of Fluid Mechanics* 45.1 (2013), pp. 409–436. ISSN: 0066-4189.

[25] Ronald J. Adrian. "Particle-Imaging Techniques for Experimental Fluid Mechanics". In: *Annual Review of Fluid Mechanics* 23.1 (1991), pp. 261–304. DOI: 10.1146/annurev.fl.23.010191.001401.

[26] A. V. Smol'yakov and V. M. Tkachenko. *The Measurement of Turbulent Fluctuations: An Introduction to Hot-Wire Anemometry and Related Transducers*. Ed. by P. Bradshaw. Berlin, Heidelberg: Springer Berlin Heidelberg, 1983, pp. 46–128. ISBN: 978-3-642-81983-4. DOI: 10.1007/978-3-642-81983-4_2. URL: https://doi.org/10.1007/978-3-642-81983-4_2.

[27] H.H Bruun. *Hot-wire anemometry : principles and signal analysis*. eng. Oxford: : Oxford University Press, 1995. ISBN: 0198563426.

[28] US National Science Foundation. Johns Hopkins Turbulence Databases. URL: http://turbulence.pha.jhu.edu/ (visited on 08/24/2018).

[29] Friedrich Pukelsheim. "The Three Sigma Rule". In: *The American Statistician* 48.2 (1994), pp. 88–91. ISSN: 0003-1305.

Appendix

Matlab scripts

PODs_and_correlations.m

```
%% INPUT

ncomphw=3; % Number of velocity components for point measurements
ncompPIV=3; % Number of velocity components for field samples
noise=0.07; % Percentage of noise added


%% LOADING DATA

load('HWdata_L704.mat'); %Single-point measurements
load('PIVdata2.mat');        %Field sample measurements


%% NOISE ADDITION

% snapshot velocities + addition of random noise
upiv=upiv+noise*randn(size(upiv));
vpiv=vpiv+noise*randn(size(vpiv));
wpiv=wpiv+noise*randn(size(upiv));

% probe velocities + addition of random noise
uhw=uhw+noise*randn(size(uhw));
vhw=vhw+noise*randn(size(vhw));
whw=whw+noise*randn(size(whw));


%% PROBE MEAN VELOCITIES
uhw_m=mean(uhw);
vhw_m=mean(vhw);
whw_m=mean(whw);
```

```
31
32
33 %% PIV MEAN VELOCITIES
34 upiv_m=mean(upiv);
35 vpiv_m=mean(vpiv);
36 wpiv_m=mean(wpiv);
37
38
39 %% PROBES' TURBULENT VELOCITIES
40 uhw_t=(uhw-uhw_m);
41 vhw_t=(vhw-vhw_m);
42 whw_t=(whw-whw_m);
43
44
45 %% PIV TURBULENT VELOCITIES
46 upiv_t=(upiv-upiv_m);
47 vpiv_t=(vpiv-vpiv_m);
48 wpiv_t=(wpiv-wpiv_m);
49
50
51 %% POD IMPLEMENTATION
52
53 %%%%%%%%%%%%%%%% Probes %%%%%%%%%%%%%%
54       if ncomphw==1
55           [psi_hw,sigma_hw,phi_hw]=svd(uhw_t(1:7000,:),'econ');
56       elseif ncomphw==2
57           [psi_hw,sigma_hw,phi_hw]=svd([uhw_t vhw_t],'econ');
58       else
59           [psi_hw,sigma_hw,phi_hw]=svd([uhw_t vhw_t whw_t],'econ');
60
61       end
62
63 phi_hw=phi_hw';
64
65 %%%%%%%%%%%%%%%% Field snapshots %%%%%%%%%%
66       if ncompPIV==2
67           [psi,sigma,phi]=svd([upiv_t vpiv_t],'econ');
68       else
69           [psi,sigma,phi]=svd([upiv_t vpiv_t wpiv_t],'econ');
70       end
71
72 phi=phi';
73
74 save('POD_PIV_','phi','psi','sigma'); %save POD from field
      snapshots
75 save('POD_HW_','phi_hw','psi_hw','sigma_hw'); %save POD from
      point measurements
```

```
76
77
78 %% CORRELATION MATRICES
79
80   size=length(psi_hw);
81   xi=(psi_hw(1:size,1:size))'*psi(1:size,1:size);
82   save('xi_','xi'); %save correlation matrix obtained
```

DataOrg_HW.m

```
%% Probe data input

up=zeros(17599, 11);
vp=zeros(17599, 11);    %saving of space for the turbulent velocity matrices
wp=zeros(17599, 11);

for i=1:17599;

    file=load(sprintf('HW_%05d.mat',i));  %loading of probe data at each time instant
    up(i,1:11)=file.UPROBE;
    vp(i,1:11)=file.VPROBE;
    wp(i,1:11)=file.WPROBE;
end

%% Computation of probe mean velocities
up_m=mean(up);
vp_m=mean(vp);
wp_m=mean(wp);

%% Obtaining of the turbulent velocity field built with the virtual probes
up_t=up-up_m;
vp_t=vp-vp_m;
wp_t=wp-wp_m;


%% Organization of the probe data for the dynamic estimation

HW=zeros(16896,23232); %saving space for the probe data matrix

for i=1:16896;

    u=reshape(up_t(i:i+703,1:11),[1,7744]);
    v=reshape(vp_t(i:i+703,1:11),[1,7744]);
    w=reshape(wp_t(i:i+703,1:11),[1,7744]);
    HW(i,:)=horzcat(u,v,w); %matrix containing the 3 turbulent velocity ccomponents
end


save('HW','HW','-v7.3') %saving the matrix containing the organized probe data
```

DynamicEstimation.m

```
%% INPUT

 ncomphw=1;
 ncompPIV=2;
 noise=7; %% percentage of noise --> Ex: 5=0.05 % noise
m=0; %% truncation rule. m=0 means no truncation is applied

%% POD's

 load(sprintf('POD_HW%dc_00%d',ncomphw,noise));    %loading the
    corresponding PODS for field and probe measurements
 load(sprintf('POD_PIV_%dc_00%d',ncompPIV,noise));

 %% CORRELATION MATRICES

 load(sprintf('xi_%d%d_00%d',ncomphw,ncompPIV,noise));  %loading
    the corresponding correlation matrix


%% Ground truth POD modes time coefficients

 load('PIVsol');

 B1=phi'*inv(sigma);
 Psi_piv=zeros(3380,7680);

 for i=1:3380;

  Psi_piv(i,:)=PIVsol(i,1:7744*ncompPIV)*B1; %the real time
     coefficients are required to compute the correlation
     coefficient

 end

 save('Psi_piv','Psi_piv');

%% TRUNCATION OF CORRELATION MATRIX

    if m>0
        stdev=(1/sqrt(7680));
        xi_t=xi;  %variable created to check the correct
           functioning of the truncation

        y=find(abs(xi)<(m*stdev));  %finding uncorrelated signals
            whose absolute value is within the truncation range
```

```
40          xi(y)=0;  %uncorrelated signal elimination
41      end
42
43  %% DYNAMIC ESTIMATION
44
45  load('HW'); %loading out-of-sample probe data for dynamic
        estimation
46  HW=HW+(noise/100)*randn(size(HW)); %addition of noise (if any)
47  range=length(sigma_hw); %control variable for matrix dimensions
48                          %(1c HW-->7000, 2c & 3c HW -->7680)
49
50  A1=phi_hw'*inv(sigma_hw);
51  A2=sigma(1:range,1:range)*phi(1:range,:);
52  Piv_Est=zeros(3380,7744*ncompPIV);
53  Psi_est=zeros(3380,range);
54  cont=0;
55
56  for i=1:5:16896;
57
58   cont=cont+1;
59   psi_HW=HW(i,1:7744*ncomphw)*A1;
60   psi_est=psi_HW*xi;
61   Psi_est(cont,:)=psi_est;
62   Piv_Est(cont,:)=psi_est*A2;
63
64  end
65
66  save('PIVest','Piv_Est'); %saving estimated field snapshots
67  save('Psi_est','Psi_est'); %saving estimated modes time
        coefficients
```

Piv_Solution.m

```
%% LOAD GROUND TRUTH DATA

upiv=zeros(3379,7744);
vpiv=zeros(3379,7744);
wpiv=zeros(3379,7744);

cont=0;
for i=1:5:16896;

    file=load(sprintf('PIV_%05d.mat',i));
    cont=cont+1;
    upiv(cont,:)=(file.UPIV)';
    vpiv(cont,:)=(file.VPIV)';
    wpiv(cont,:)=(file.WPIV)';

end

%mean velocities
upiv_m=mean(upiv);
vpiv_m=mean(vpiv);
wpiv_m=mean(wpiv);

%turbulent velocities
upiv_t=upiv-upiv_m;
vpiv_t=vpiv-vpiv_m;
wpiv_t=wpiv-wpiv_m;

PIVsol=horzcat(upiv_t,vpiv_t,wpiv_t); %ground truth PIV matrix
```

Parametric_Study.m

```
%% INPUT

ncomphw=3; %Number of components for probe measurements
ncompPIV=3; %Number of components for field samples
m=2;  % Filtering rule factor. m=0 for non filtered case

x=linspace(0,1,88);
y=x;
%% Squared correlation coefficient R^2

if m==0

    load(sprintf('Psi_est%d%d',ncomphw,ncompPIV));

else
    load(sprintf('Psi_est%d%d_m%d',ncomphw,ncompPIV,m));
end

load(sprintf('Psi_piv%d',ncompPIV));

size=length(Psi_est);
Psi_piv=Psi_piv(:,1:size);
C1=zeros(3380,size);

for i=1:3380

    C1(i,:)=Psi_est(i,:).*(Psi_piv(i,:));

end

C2=(mean(C1)).^2;
C3=mean((Psi_est).^2).*mean((Psi_piv).^2);

Rcorr=C2./C3;
RcorrS=smooth(Rcorr); %smoothing of the resulting curve


%% Squared determination coefficient D2

det_e2=(Psi_piv(:,1:length(Psi_est))-Psi_est).^2;
SSres=sum(det_e2);
SStot=sum((Psi_piv-mean(Psi_piv)).^2);


Rdet=1-(SSres./SStot(:,1:length(Psi_est)));
```

```
46 RdetS=smooth(Rdet);  %smoothing of the resulting curve
47
48
49 % plot both coefficients
50 figure
51 grid on
52 semilogx(RcorrS(1:150),'*','Linewidth',1,'Markersize',6);
53 xlabel('$\textbf{Modes}$','interpreter','latex','fontsize',20)
54 ylabel('$\textbf{Rcorr, Rdet}$','interpreter','latex','fontsize'
       ,20);
55 set(gca,'TickLabelInterpreter','latex','fontsize',20)
56
57 hold on
58 semilogx(RdetS(1:150),'-','Linewidth',1.5);
59 title('$\textbf{Correlation and determination coefficients}$','
      interpreter','latex','fontsize',22);
60 h=legend('Rcorr','Rdet');
61 set(h,'Location','Northeast','interpreter','latex','fontsize',20)
62
63
64 %% Velocity correlation
65
66 load('PIVsol');
67
68 if m==0
69
70     load(sprintf('PIVest%d%d',ncomphw,ncompPIV));
71
72 else
73     load(sprintf('PIVest%d%d_m%d',ncomphw,ncompPIV,m));
74 end
75
76
77 u_real=PIVsol(:,1:7744);
78 u_est=Piv_Est(:,1:7744);
79
80 st_real=std(u_real);
81 st_est=std(u_est);
82
83 n=zeros(3380,7744);
84
85 for i=1:3380
86
87    n(i,:)=(u_est(i,:).*u_real(i,:));
88
89 end
90
```

```
91  point_corr=(mean(n)./(st_est.*st_real)).*100;
92
93
94  %plot
95  figure
96  imagesc(x,y,reshape(point_corr,[88,88]));
97  title('$\textbf{Velocity correlation}$','interpreter','latex','
        fontsize',22);
98  xlabel('$x/h$','interpreter','latex','fontsize',20)
99  ylabel('$y/h$','interpreter','latex','fontsize',20)
100 h=colorbar;
101 set(h,'fontsize',20,'TickLabelInterpreter','latex')
102 h.Label.Interpreter = 'latex';
103 h.Label.String = '$[\%]$';
104 set(gca,'Ydir','Normal','TickLabelInterpreter','latex','fontsize'
        ,20)
105
106
107 %% Error estimation
108
109 U_bulk=1; %DNS bulk velocity used for adimensionalization of
        results
110 e=(sqrt(sum((u_est-u_real).^2))./(3380))./U_bulk;
111 error=(reshape(e,[88,88])).*100; %estimation error in [%]
112
113 figure
114 pcolor(x,y,error);
115 shading interp;
116 colormap jet;
117 title('Average estimation error','interpreter','latex','fontsize'
        ,30);
118 xlabel('$x/h$','interpreter','latex','fontsize',18);
119 ylabel('$y/h$','interpreter','latex','fontsize',18);
120 set(gca,'Ydir','Normal','TickLabelInterpreter','latex','fontsize'
        ,18);
121 lim2=caxis;
122 pbaspect([1 1 1]);
123 c=colorbar;
124 set(c,'fontsize',18,'TickLabelInterpreter','latex');
125 c.Label.Interpreter='latex';
126 c.Label.String = '$\overline{Err}/U_b$ $\left[\%\right]$';
127 c.Label.Rotation = 0;
128 set(c,'YTick',0:0.05:0.25)
```